

GenCore version 5.1.7  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - nucleic search, using **lsame.plue.p2n.model**

Run on: February 17, 2006, 14:34:01 ; Search time 7245 seconds  
(without alignments)  
3601.262 Million cell updates/sec

Title: US-09-967-237B-2  
Perfect score: 2424  
Sequence: 1 MAPLCPSPWLLPAPAPG.....RRGKTKGVSYRPAEVAETGA 459

Scoring table: BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 5883141 seqs, 28421725653 residues  
Total number of hits satisfying chosen parameters: 11766282

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters: -MODEL=frame.p2n.model -DRV=xlh  
-Q=/abs/ABSSWEB.spool/US09967237/runat\_15022006\_182242\_27965/app\_query.fasta\_1  
-DB=GenEmbl -OPMT=fastap -SUFFIX=rge -MINMATCH=0.1 -LOOPL=0 -LOOPEXT=0  
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45  
-DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTPMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000 -HOST=abs03h  
-USER=US09967237@cgn\_1\_4939@runat\_15022006\_182242\_27965 -NCPU=6 -ICPU=3  
-NO MMAP -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG -DEV TIMEOUT=120  
-WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7  
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : GenEmbl.\*  
1: gb\_ba.\*  
2: gb\_in.\*  
3: gb\_env.\*  
4: gb\_cm.\*  
5: gb\_ov.\*  
6: gb\_pat.\*  
7: gb\_ph.\*  
8: gb\_pr.\*  
9: gb\_ro.\*  
10: gb\_sts.\*  
11: gb\_sy.\*  
12: gb\_un.\*  
13: gb\_vi.\*  
14: gb\_htg.\*  
15: gb\_pl.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES			
Result No.	Score	Query Match Length	Description
1	2424	100.0	1519 8 HAJ10588 Homo sapi
2	2424	100.0	1522 6 AR074439 Sequence
3	2424	100.0	1522 6 AR081119 Sequence

ALIGNMENTS			
RESULT 1	HAJ10588	1519 bp	linear PRI 15-APR-2005
LOCUS	Homo sapiens mRNA for renal cell carcinoma associated antigen G250.		
DEFINITION	AJ010588		
ACCESSION	AJ010588.1	GI:7327887	
VERSION	Renal cell carcinoma associated antigen G250.		
KEYWORDS	Renal cell carcinoma associated antigen G250		
SOURCE	Homo sapiens		
ORGANISM	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
REFERENCE	1		
AUTHORS	Grabmaier, K., Visser, J.L., De Weijert, M.C., Oosterwijk-Wakka, J.C., Van Bokhoven, A., Brakenhoff, R.H., Noessner, E., Mulders, P.A., Merks, G., Figdor, C.G., Adema, G.J. and Oosterwijk, E.		
TITLE	Molecular cloning and immunogenicity of renal cell carcinoma-associated antigen G250		
JOURNAL	Int. J. Cancer 85 (6), 865-870 (2000)		
PURNED	10709109		
REFERENCE	2 (bases 1 to 1519)		
AUTHORS	Oosterwijk, E.		
TITLE	Direct Submission		
JOURNAL	Submitted (26-AUG-1998) Oosterwijk E., St. Radboud Academic Hospital Nijmegen, Urological Research Laboratory, Postbus 9101, 6500 HB Nijmegen, THE NETHERLANDS		
FEATURES	Location/Qualifiers		

source	1. .1519 /organism="Homo sapiens" /mol_type="mRNA" /db_xref="taxon:9606" /chromosome="9" /map="p12-13" /cell_line="SK-RC-10" /tissue_type="renal cell carcinoma" <1. .9 10. .1389 /codon_start=1 /product="renal cell carcinoma associated antigen G250" /protein_id="CA82444.1" /db_xref="GI:7327888" /db_xref="GOA:Q16790" /db_xref="UniProt:IPR001148" /trna_labels="MAPLCPSPMLPILIPAPGLTVQLLSILLMPVHPORLPBQ EDSPGGSGEDDPLGEDLPSEEDSPREDDPPGEDLPGEEDLPGEEDLPVKPKS EEGSDKLEDPTEAPDPQEPONNAHRDKGDDQSHRWYGGDPPWPRVSPACGRF QSEVDIRPOLAACPALRLELGLFQLPELPELRNNGHSVOLTPPGLMALGPR EYRALQLHFWGAAGRPGEHTVEGRFAEIHVHLSTAFARVDFBALGRPGGLAVIA AFUEGPEENSAVEQLLSLEEIAEGSSTQVPGLDI SALLPSDFSRXYQYEGSLTTP PCAQGVITWVFNQTVMLSAKQLHTLSDTLWPGDSRLQNFRTATQPLNGRVLEASPPA GVDSFRAAEFVQLNSCLAAGDILALVFGLLFAVTSVAFIVQMRHRRGTGKGVSYR PAEVAETGA"
5'UTR	10. .120
CDS	412. .1179 /product="carbonic anhydrase" 106 /note="polymorphism" /replaces="g" 1252. .1311 /notes="transmembrane region" 1390. .1519 1504. .1509
sig_peptide	
mat_peptide	
variation	
misc_feature	
3'UTR	
polyA_signal	
ORIGIN	
Alignment Scores:	
Pred. No.:	6.19e-103 Length: 1519
Score:	2424.00 Matches: 459
Percent Similarity:	100.0% Conservative: 0
Best Local Similarity:	100.0% Mismatches: 0
Query Match:	100.0% Indels: 0
DB:	8 Gaps: 0
US-09-967-237B-2 (1-459) x HAU10588 (1-1519)	
QY	1 MetAlaProLeuCySProSerProTTPLeuProLeuLeuLeuLeuProAlaProGly 20
Db	10 ATGGCTCCCTGTGCCCCAGCCCTGGCTCCCTCTGTGATCCCGGCCCTGTCTCAGGC 69
QY	21 LeuThrValGlnLeuLeuLeuSerLeuLeuLeuLeuMetProValHisProGlnArgLeu 40
Db	70 CTCACTGTGCAACTGTCTGTCTCACTGCTCTCTCTGATGCTCTCCATCCCCAGAGTTG 129
QY	41 ProArgMetGlnGluAspSerProLeuGlyGlyGlySerGlyGluAspProLeu 60
Db	130 CCCCGATGTCAGAGGATTCCTCCCTGGGAGGAGGCTCTTCTGGGGAAGATGACCCACTG 189
QY	61 GlyGluGluAspLeuProSerGluGluAspSerProArgGluGluAspProProGlyGlu 80
Db	190 GCGGAGGAGGATTCGCCAGTGAAGAGGATTCACCCAGAGAGGAGATCCACCCGAGAG 249
QY	81 GluAspLeuProGlyGluGluAspLeuProGlyGluGluAspLeuProGluValLysPro 100
Db	250 GAGGATCTACTGGAGAGGAGGATCTACCTGGAGAGGAGGATCTACTGAAGTTAAGCCT 309
QY	101 LysSerGluGlnGluGlySerLeuLysLeuGluAspLeuProThrValGluAlaProGly 120
Db	310 AAATCAAGAGAGAGGCTCCCTGAAAGTTAGAGGATCTACTACTGTTGAGGCTCTCTGA 369
QY	121 AspProGlnGluProGlnAsnAlaHisArgAspLysGluGlyAspGlnSerHis 140

RESULT 2  
AR074439  
LOCUS  
Sequence 1 from patent US 5955075.  
AR074439  
ACCESSION  
VERSION  
AR074439.1 GI:10001194

AR074439 1522 bp DNA linear PAT 28-AUG-2000  
Sequence 1 from patent US 5955075.  
AR074439  
ACCESSION  
VERSION  
AR074439.1 GI:10001194

|||||  
370 GATCTCTCAAGAACCCAGAAATAATGCCACAGGAGCAAGAGGGATGACACAGATCAT 429  
|||||  
141 TTPATGTyrGlyGlyAspProProTTPProArgValSerProAlaCysAlaGlyArgPhe 160  
|||||  
430 TGGCGCTATGGAGGCGACCCGCCCTGGCCCCGGGTGTCCCCAGCCTGCGCGGCGCCTTC 489  
|||||  
161 GlnSerProValAspIleArgProGlnLeuAlaAaPheCysProAlaLeuArgProLeu 180  
|||||  
490 CAGTCCCGCGTGGATATCCGCCCCAGCTCGCGCCCTTCTGCGCGCCCTCGCGCCCTCG 549  
|||||  
181 GluLeuLeuGlyPheGlnLeuProProLeuProGluLeuArgLeuArgAsnGlnHis 200  
|||||  
550 GAACCTCTGGGCTTCAGCTCCGCCGCTCCAGAACTCGCGCTCGCGCAACAATATGCCAC 609  
|||||  
201 SerValGlnLeuThrLeuProProGlyLeuGluMetAlaLeuGlyProGlyArgGluTyr 220  
|||||  
610 AGTGTGCAACTGACCCCTGCTCTCGGCTAGAGATGGCTCTGGGTCCCGGGGGAGTAC 669  
|||||  
221 ArgAlaLeuGlnLeuHisLeuHisTTPGlyAlaAaGlyArgProGlySerGluHisThr 240  
|||||  
670 CGGGCTCTGCAGTGCATCTGCACCTGGGGGGCTGCAGGTCGTCCGGGCTCGGAGCACCT 729  
|||||  
241 ValGluGlyHisArgPheProAlaGluIleHisValValHisLeuSerThrAlaPheAla 260  
|||||  
730 GTGGAAAGGCCCGCTTCCCTGCCGAGATCCACGTGGTTCACCTCAGCACCCGCTTTGCC 789  
|||||  
261 ArgValAspGluAlaLeuGlyArgProGlyGlyLeuAlaValLeuAlaAaPheLeuGlu 280  
|||||  
790 AGAGTTGACGAGGCTTGGGGCGCCCGGAGGCTGGCGGTGTGGCCGCTTTCTGGAG 849  
|||||  
281 GluGlyProGluGluAsnSerAlaTyrGluGlnLeuLeuSerArgLeuGluIleAla 300  
|||||  
850 GAGGGCCCGGAGAAACACAGTGCCTATGACAGTGTGCTGTCTCGCTTGGAAAGAAATCGCT 909  
|||||  
301 GluGluGlySerGluThrGlnValProGlyLeuAspIleSerAlaLeuLeuProSerAsp 320  
|||||  
910 GAGGAAGGCTCAGAGACTCAGGTCCAGGACTGGACATATCTGCATCTCTCGCTCTGAC 969  
|||||  
321 PheSerArgTyrPheGlnTyrGluGlySerLeuThrThrProProCysAlaGlnGlyVal 340  
|||||  
970 TTCAGCGGCTACTTTCCTCAATATGAGGGGTCTCTGACTACACCGCCCTGTGCCAGGGGTGC 1029  
|||||  
341 IleTTPThrValPheAsnGlnThrValMetLeuSerAlaGlyGlnLeuHisThrLeuSer 360  
|||||  
1030 ATCTGAGCTGTGTAAACAGACAGTGTGCTGAGTGTGCTGAGGCTCCACACCCCTCTCT 1089  
|||||  
361 AspThrLeuTTPGlyProGlyAspSerArgLeuGlnLeuAsnPheArgAlaThrGlnPro 380  
|||||  
1090 GACACCTGTGGGACCTGTGTGACTCTCGGCTACAGTGAACCTTCGAGGGAGCGACGCT 1149  
|||||  
381 LeuAsnGlyArgValIleGluAlaSerPheProAlaGlyValAspSerSerProArgAla 400  
|||||  
1150 TTTGAATGGCGAGTGATTGAGGCTCTTCCCTGCTGGAGTGGACAGCAGTCTCTCGGGCT 1209  
|||||  
401 AlaGluProValGlnLeuAsnSerCysLeuAlaAaGlyAspIleLeuAlaLeuValPhe 420  
|||||  
1210 GCTGAGCCAGTCCAGCTGAATTCCTGCTGGCTGGTGGTGGACATCTCTAGCCCTGGTTTT 1269  
|||||  
421 GlyLeuLeuPheAlaValThrSerValAlaPheLeuValGlnMetArgArgGlnHisArg 440  
|||||  
1270 GGCCTCTTTTGTGTCTACAGGCTCGGCTGCTCTTGTGCGAGTGAAGAGGAGCAGCAGA 1329  
|||||  
441 ArgGlyThrLysGlyGlyValSerTyrArgProAlaGluValAlaGluThrGlyAla 459  
|||||  
1330 AGGGGAACCAAGGGGGTGTGAGTACCAGCCAGCAGAGGTAGCCGAGACTGGAGCC 1386  
|||||



```
QY 1 MetAlaProLeuCysProSerProTrrpLeuProLeuLeuIleProAlaProAlaProGly 20
Db 13 ATGGCTCCCTGTGCCCCAGCCCTGGCTCCCTCTGTGTGATCCCGGCCCTGTCTCCAGGC 72
QY 21 LeuThrValGlnLeuLeuLeuSerLeuLeuLeuLeuMetProValHisProGlnArgLeu 40
Db 73 CTCACTGTGCACTGTCTGTCTGCTGCTCTGTGATGCTGTCTCAATCCCGCAGAGGTG 132
QY 41 ProArgMetGlnGluAspSerProLeuGlyGlySerSerGlyGluAspAspProLeu 60
Db 133 CCCCGATGCAGAGGATCCCTTGGGAGGAGGCTCTTCTGGGAAAGATGACCCACTG 192
QY 61 GlyGluGluAspLeuProSerGluGluAspSerProArgGluGluAspProProGlyGlu 80
Db 193 GCGAGAGGATCTGCCCTGAAGAGGATTCACCCAGAGAGGAGATCCACCCGAGAG 252
QY 81 GluAspLeuProGlyGluGluAspLeuProGlyGluGluAspLeuProGluValHisPro 100
Db 253 GAGGATCTACCTGGAGAGGAGGATCTACCTGGAGAGGAGGATCTACCTGAAGTAAAGCCT 312
QY 101 LysSerGluGluGluGlySerLeuHisLeuGluAspLeuProThrValGluAlaProGly 120
Db 1033 ATCTGACGTGTGTAAACAGAGAGTGTGCTGAGTGTGCTAAGCAGCTCCACACCTCTCT 1092
QY 361 AspThrLeuTrpGlyProGlyAspSerArgLeuGlnLeuAsnPheArgAlaThrGlnPro 380
```

```
Db 1093 GACACCTGTGGGACCTGTGTGACTCTCGCTACAGCTGAACCTTCGAGGAGCGACGCT 1152
QY 381 LeuAsnGlyArgValIleGluAlaSerPheProAlaGlyValAspSerSerProArgAla 400
Db 1153 TTGAATGGCGGAGTGATTCAGGCTCTCTTCCCTGCTGGAGTGGACAGCAGTCTCGGGCT 1212
QY 401 AlaGluProValGlnLeuAsnSerCysLeuAlaGlyAspIleLeuAlaLeuValPhe 420
Db 1213 GCTGAGCCAGTCCAGCTGAATTCCTGCTGGCTGTGCTGGTGACATCTAGCCCTGGTCTT 1272
QY 421 GlyLeuLeuPheAlaValThrSerValAlaPheLeuValGlnMetArgArgGlnHisArg 440
Db 1273 GGCCTCTCTTTTGTCTGTACAGCGTCCGCTTCTTGTGCAGTGAAGAGGAGCAGACA 1332
QY 441 ArgGlyThrLysGlyGlyValSerTyrArgProAlaGluValAlaGluThrGlyAla 459
Db 1333 AGGGGAACCAAGGGGGTGTGAGCTACCGCCAGCAGAGGTAGCCGAGACTGGAGCC 1389

RESULT 4
AR085316
LOCUS 1522 bp DNA linear PAT 01-SBP-2000
DEFINITION Sequence 1 from patent US 5981711.
ACCESSION AR085316
VERSION AR085316.1 GI:10012085
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 1522)
AUTHORS Zavada,J., Pastorekova,S. and Pastorek,J.
TITLE MN-specific antibodies and hybridomas
JOURNAL Patent: US 5981711-A 1 09-NOV-1999;
FEATURES
source
1..1522
/organism="unknown"
/mol_type="unassigned DNA"

ORIGIN

Alignment Scores:
Pred. No.: 6.2e-103 Length: 1522
Score: 2424.00 Matches: 459
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 6 Gaps: 0

US-09-967-237B-2 (1-459) x AR085316 (1-1522)
QY 1 MetAlaProLeuCysProSerProTrrpLeuProLeuLeuIleProAlaProAlaProGly 20
Db 13 ATGGCTCCCTGTGCCCCAGCCCTGGCTCCCTCTGTGTGATCCCGGCCCTGTCTCCAGGC 72
QY 21 LeuThrValGlnLeuLeuSerLeuLeuLeuLeuMetProValHisProGlnArgLeu 40
Db 73 CTCACTGTGCAACTGTCTGTCTGCTGCTTCTGATGCTTCTGATGCCCTCCAGAGGTG 132
QY 41 ProArgMetGlnGluAspSerProLeuGlyGlySerSerGlyGluAspAspProLeu 60
Db 133 CCCCGATGCAGAGGATTCCTCCCTTGGGAGGAGGCTCTTCTGGGAAAGATGACCCACTG 192
QY 61 GlyGluGluAspLeuProSerGluGluAspSerProArgGluGluAspProProGlyGlu 80
Db 193 GCGAGAGGATCTGCCCTGAAGAGGATTCACCCAGAGAGGAGATTCACCCGAGAG 252
QY 81 GluAspLeuProGlyGluGluAspLeuProGlyGluGluAspLeuProGluValHisPro 100
Db 253 GAGGATCTACCTGGAGAGGAGGATCTACCTGGAGAGGAGGATCTACCTGAAGTAAAGCCT 312
QY 101 LysSerGluGluGluGlySerLeuHisLeuGluAspLeuProThrValGluAlaProGly 120
Db 313 AAATCAGAGAGAGGGCTCCCTGAAAGTTAGAGGATCTACTACTGTGAGGCTCCCTGGA 372
```



QY	241	ValGluGlyHisArgPheProAlaGluIleHisValHisLeuSerThrAlaPheAla	260
Db	733	GTGGAGGCCACCGTTTCCCTGCGGATCCAGTGTTCACTCAGACCGCTTTGCC	792
QY	261	ArgValAspGluAlaLeuGlyArgProGlyGlyLeuAlaValLeuAlaPheLeuGlu	280
Db	793	AGAGTTGACGAGGCTTGGGGGCCCGGGAGGCTGGCGTGTGGCGGCTTTCTGGAG	852
QY	281	GluGlyProGluGluAsnSerAlaTyrGluGlnLeuSerArgLeuGluIleAla	300
Db	853	GAGGGCCCGGAGAAACAGTGCTATGACAGTGTGCTGCTGGGAAGAAATCGCT	912
QY	301	GluGluGlySerGluThrGlnValProGlyLeuAspIleSerAlaLeuLeuProSerAsp	320
Db	913	GAGGAGGCTCAGAGACTCAGGTCCAGGACTGGACATATCTGCATCTCCCTCTGAC	972
QY	321	PheSerArgTyrPheGlnTyrGluGlySerLeuThrThrProCysAlaGlnGlyVal	340
Db	973	TTGAGCGGCTACTTCCAATATGAGGGTCTCTGACTACACCGCCCTGTGCCAGGGTGTG	1032
QY	341	IleTyrThrValPheAsnGlnThrValMetLeuSerAlaLysGlnLeuHisThrLeuSer	360
Db	1033	ATCTGGACTGTGTTAAACACAGACAGTGTCTGAGTGCTAAGCAGCTCCACACCTCTCT	1092
QY	361	AspThrLeuTyrGlyProGlyAspSerArgLeuGlnLeuAsnPheArgAlaThrGlnPro	380
Db	1093	GACACCTGTGGGACCTGTGACTCTGGCTACAGCTGAACCTTCGAGGCGACGACCT	1152
QY	381	LeuAsnGlyArgValIleGluAlaSerPheProAlaGlyValAspSerProArgAla	400
Db	1153	TTCAATGGGCGAGTGATTGAGGCTCTCTCCCTGCTGGAGTGACAGAGCTCTCGGGCT	1212
QY	401	AlaGluProValGlnLeuAsnSerCysLeuAlaGlyAspIleLeuAlaLeuValPhe	420
Db	1213	GCTGAGCCAGTCCAGCTGAATCTGCTGCTGGCTGCTGGTGACATCCTAGCCCTGGTTTTT	1272
QY	421	GlyLeuLeuPheAlaValThrSerValAlaPheLeuValGlnMetArgArgGlnHisArg	440
Db	1273	GGCTCTCTTTTGTGTACACAGCTGCGCTTCTTGTGAGATGAGAGGCGACAGACA	1332
QY	441	ArgGlyThrIysGlyGlyValSerTyrArgProAlaGluValAlaGluThrGlyAla	459
Db	1333	AGGGGAACCAAGGGGGTGTGAGCTACCGCCAGAGAGGTAGCCGAGACTGGAGCC	1389
RESULT 6			
AR104223			
LOCUS			
DEFINITION			
ACCESSION			
VERSION			
KEYWORDS			
SOURCE			
ORGANISM			
REFERENCE			
1 (bases 1 to 1522)			
AUTHORS			
Zavada,J., Pastorekova,S. and Pastorek,J.			
TITLE			
Detection and quantitation of MN-specific antibodies			
JOURNAL			
Patent: US 6093548-A 1 25-JUL-2000;			
FEATURES			
Location/Qualifiers			
1..1522			
source			
/organism="unknown"			
/mol_type="unassigned DNA"			
ORIGIN			
Alignment Scores:			
Pred. No.:			
Score:			
Percent Similarity:			
Best Local Similarity:			
Query Match:			
DB:			
6..2e+103			
2424.00			
100.0%			
100.0%			
100.0%			
Length:			
Matches:			
Conservative:			
Mismatch:			
Indels:			
Gaps:			
1522			
459			
0			
0			
0			
6			
US-09-967-237B-2 (1-459) x AR104223 (1-1522)			

QY	1	MetAlaProLeuCysProSerProTyrLeuProLeuLeuIleProAlaProGly	20
Db	13	ATGGCTCCCTGTGCCCCCAGCCCTTCTGTCTGTTGATCCCGCCCTTGTCTCAGGC	72
QY	21	LeuThrValGlnLeuLeuSerLeuLeuLeuMetProValHisProGlnArgLeu	40
Db	73	CTCAGTGTGCACTGTGTGTCTGCTGCTTCTGATGCTTCTCATCCCCAGAGGTG	132
QY	41	ProArgMetGlnGluAspSerProLeuGlyGlySerSerGlyGluAspAspProLeu	60
Db	133	CCCCGATGACAGAGGATTCCTCCCTTGGGAGGAGGCTCTTCTGGGGAAGATGACCCACTG	192
QY	61	GlyGluGluAspLeuProSerGluGluAspSerProArgGluGluAspProProGlyGlu	80
Db	193	GGCGAGAGGATCTGCCAGTGAAGAGGATTCACCCAGAGAGGAGATCCACCCGAGAG	252
QY	81	GluAspLeuProGlyGluGluAspLeuProGlyGluGluAspLeuProGluValLysPro	100
Db	253	GAGGATCTACCTGGAGAGGAGGATCTACCTGGAGAGGAGGATCTACTGAAGTTAGCCT	312
QY	101	LysSerGluGluGlySerLeuLysLeuGluAspLeuProThrValGluAlaProGly	120
Db	313	AAATCAGAAAGAGGCTCCCTGAAGTTAGAGGATCTACTGTTGAGGCTCTCTGGA	372
QY	121	AspProGlnGluProGlnAsnAsnAlaHisArgAspLysGluGlyAspAspGlnSerHis	140
Db	373	GATCTTCAGAAACCCAGAAATATGCCACAGGACAAAGAGGAGGATGACCAAGTGCAT	432
QY	141	TyrArgTyrGlyGlyAspProProTyrProArgValSerProAlaCysAlaGlyArgPhe	160
Db	433	TGGCGCTATGAGGAGGACCCGCTTGGCCCGGGTGTCCCGAGCTTGGGGGCGCTTC	492
QY	161	GlnSerProValAspIleArgProGlnLeuAlaAlaPheCysProAlaLeuArgProLeu	180
Db	493	CAGTCCCGGTGGATATCCGCTCCGCTCCGCTTCTGCTCCGCTTGGGGGCGCTTC	552
QY	181	GluLeuLeuGlyPheGlnLeuProProLeuProGluLeuArgLeuArgAsnGlyHis	200
Db	553	GAACTCTCGGGCTTCAGCTCCGCTCCGCTCCAGAACTCGGCTCGGCAACAAATGGCCAC	612
QY	201	SerValGlnLeuThrLeuProProGlyLeuGluMetAlaLeuGlyProGlyArgGluTyr	220
Db	613	AGTGTGCACTGACCTGCTCTCTGGCTTAGAGATGGCTCTGGGTCCCGGGGAGTAC	672
QY	221	ArgAlaLeuGlnLeuHisLeuHisTyrGlyAlaAlaGlyArgProGlySerGluHisThr	240
Db	673	CGGGCTCTGAGCTGCATCTGCAGTGGGGGCTGCGAGTCTGTCGGGCTCGGAGCACACT	732
QY	241	ValGluGlyHisArgPheProAlaGluIleHisValValHisSerThrAlaPheAla	260
Db	733	GTGGAGGCGCACCGTTTCTCCGAGATCCACGTTGTTTCTGCTGCTGCTTGGCTTGGC	792
QY	261	ArgValAspGluAlaLeuGlyArgProGlyGlyLeuAlaValLeuAlaPheLeuGlu	280
Db	793	AGAGTTGACAGGCTTGGGGGCGCGGAGGCGCTGGCGCTGTGGCGGCTTCTGGAG	852
QY	281	GluGlyProGluGluAsnSerAlaTyrGluGlnLeuLeuSerArgLeuGluIleAla	300
Db	853	GAGGGCCCGGAAGAAACAGTGCCTATGAGCAGTGTCTGCTGCTTGGAAAGAAATCGCT	912
QY	301	GluGluGlySerGluThrGlnValProGlyLeuAspIleSerAlaLeuLeuProSerAsp	320
Db	913	GAGGAGGCTCAGAGACTCAGGTCCAGAGCTGGACATATCTGCACTCTCTGCTCTGAC	972
QY	321	PheSerArgTyrPheGlnTyrGluGlySerLeuThrThrProCysAlaGlnGlyVal	340
Db	973	TTGAGCGGCTACTTCCAATATGAGGGTCTCTGACTACACCGCCCTGTCGCCAGGGTGC	1032
QY	341	IleTyrThrValPheAsnGlnThrValMetLeuSerAlaLysGlnLeuHisThrLeuSer	360
Db	1033	ATCTGGACTGTGTTAAACACAGACAGTGTCTGAGTGCTAAGCAGCTCCACACCTCTCT	1092

```
QY 361 AspThrLeuTyrGlyProGlyAspSerArgLeuGlnLeuAsnPheArgAlaThrGlnPro 380
|
|
|
Db 1093 GACACCTGTGGGACCTGTGACTCTCGCTACAGCTGACTTCCGAGCGACGACCT 1152
|
|
|
QY 381 LeuAsnGlyArgValIleGluAlaSerPheProAlaGlyValAspSerProArgAla 400
|
|
|
Db 1153 TTGAATGGGCGAGTGTAGGCGCTCTTCCCTGCTGGAGTGACAGCAGTCTCGGGCT 1212
|
|
|
QY 401 AlaGluProValGlnLeuAsnSerCysLeuAlaAlaGlyAspIleLeuAlaLeuValPhe 420
|
|
|
Db 1213 GCTGAGCCAGTCCAGCTGAATTCCTGCTGGCTGCTGGTGACATCTAGCCCTGGTTTTT 1272
|
|
|
QY 421 GlyLeuLeuPheAlaValThrSerValAlaPheLeuValGlnMetArgArgGlnHisArg 440
|
|
|
Db 1273 GGCTCTCTTTTGTGTCACCGGCTCGCTTCTTGTGACATGAGAGGCGACGACAGA 1332
|
|
|
QY 441 ArgGlyThrIysGlyGlyValSerTyrArgProAlaGluValAlaGluThrGlyAla 459
|
|
|
Db 1333 AGGGGAACCAAGGGGTGTGAGCTACCGCCAGCAGAGTAGCCGAGACTGGAGCC 1389

RESULT 7
LOCUS AR143487 1522 bp DNA linear PAT 08-AUG-2001
DEFINITION Sequence 1 from patent US 6204370.
ACCESSION AR143487
VERSION AR143487.1 GI:15104773
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 1522)
AUTHORS Zavada,J., Pastorekova,S. and Pastorek,J.
TITL MN gene and protein
JOURN Patent: US 6204370-A 1 20-MAR-2001;
FEATURES Location/Qualifiers
source 1..1522
/organism="unknown"
/mol_type="unassigned DNA"

ORIGIN

Alignment Scores:
Pred. No.: 6 2e-103 Length: 1522
Score: 2424.00 Matches: 459
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 6 Gaps: 0

US-09-967-237B-2 (1-459) x AR143487 (1-1522)

QY 1 MetAlaProLeuCysProSerProTyrLeuProLeuLeuIleProAlaProGly 20
|
|
|
Db 13 ATGGCTCCCTGTGCCCCAGCCCTGGCTCCCTCTGTGTGATCCCGGCCCTGTCTCAGGC 72
|
|
|
QY 21 LeuThrValGlnLeuLeuLeuSerLeuLeuLeuMetProValHisProGlnArgLeu 40
|
|
|
Db 73 CTCACTGTGCACTGTGCTGTCTCACTGCTCTTCTGATGCTCTCCATCCCGAGAGTTG 132
|
|
|
QY 41 ProArgMetGlnGluAspSerProLeuGlyGlySerSerGlyGluAspProLeu 60
|
|
|
Db 133 CCCCCGATGAGGAGGATCCCCCTTGGGAGGAGGCTCTTCTGGGGAAGATGATGCCACTG 192
|
|
|
QY 61 GlyGluGluAspLeuProSerGluGluAspSerProArgGluAspProGlyGlu 80
|
|
|
Db 193 GCGAGAGAGATTCGCCAGTAGAGGATTCACCCAGAGAGAGGATCCACCCGAGAG 252
|
|
|
QY 81 GluAspLeuProGlyGluGluAspLeuProGlyGluAspLeuProGluValLysPro 100
|
|
|
Db 253 GAGATCTACTGTGAGGAGGAGGATCTACCTGTGAGAGGAGGATCTACTGAAGTTAGCCT 312
|
|
|
QY 101 LysSerGluGluGluGlySerLeuLeuLeuGluAspLeuProThrValGluAlaProGly 120
|
|
|
Db 313 AAATCAGAGAGAGGGCTCCCTGAAGTTAGAGGATCTACTACTGTGTGAGGCTCTCTGGA 372
```

```
QY 121 AspProGlnGluProGlnAsnAlaHisArgAspLysGluGlyAspAspGlnSerHis 140
|
|
|
Db 373 GATCTCTCAGAACCCAGATATATGCCACAGGACAAAGAGGGGATGACACAGTCTAT 432
|
|
|
QY 141 TrpArgTyrGlyGlyAspProProTyrProArgValSerProAlaCysAlaGlyArgPhe 160
|
|
|
Db 433 TGGCGCTATGAGGAGGACCCGCCCTGGCCCGGGTGTCCCGACCTGCGCGCGCGCTTC 492
|
|
|
QY 161 GlnSerProValAspIleArgProGlnLeuAlaAlaPheCysProAlaLeuArgProLeu 180
|
|
|
Db 493 CAGTCCCCCGTGGATATCCGCCCCAGCTCGCCGCCCTTCTGCCGCCCTTGGCCCCCTG 552
|
|
|
QY 181 GluLeuLeuGlyPheGlnLeuProProLeuProGluLeuArgLeuArgAsnGlyHis 200
|
|
|
Db 553 GAACCTCTGGCTTCCAGCTCCCGCGCTCCAGAACTGCGGCTCGCAACAATGSCCAG 612
|
|
|
QY 201 SerValGlnLeuThrLeuProProGlyLeuGluMetAlaLeuGlyProGlyArgGluTyr 220
|
|
|
Db 613 AGTGTGCAACTGACCTGCTCTGGCTTAGAGATGGCTCTGGGTCCCGCGCGGAGTAC 672
|
|
|
QY 221 ArgAlaLeuGlnLeuHisLeuHisTyrGlyAlaAlaGlyArgProGlySerGluHisThr 240
|
|
|
Db 673 CGGGCTCTCGAGCTGCATCTGCACCTGGGGGGCTGCGAGTCTGTCGGGCTCGGAGCAC 732
|
|
|
QY 241 ValGluGlyHisArgPheProAlaGluIleHisValHisLeuSerThrAlaPheAla 260
|
|
|
Db 733 GTGGAGGCGCACCGTTCCTGCGGAGATCCAGCTGTTCACTCAGCAGCCGCTTTGCC 792
|
|
|
QY 261 ArgValAspGluAlaLeuGlyArgProGlyGlyLeuAlaValLeuAlaAlaPheLeuGlu 280
|
|
|
Db 793 AGAGTTGACAGGCGCTTGGGGGCGCCCGGAGGCGCTGCGCGTGTGTCGCGCTTTCTGGAG 852
|
|
|
QY 281 GluGlyProGluGluAsnSerAlaTyrGluGlnLeuLeuSerArgLeuGluIleAla 300
|
|
|
Db 853 GAGGGCCCGAAGAAACAGTGCCTATGAGCAGTTCCTGCTGCTTGAAGAAATTCGCT 912
|
|
|
QY 301 GluGluGlySerGluThrGlnValProGlyLeuAspIleSerAlaLeuLeuProSerAsp 320
|
|
|
Db 913 GAGGAAGGCTCAGAGACTCAGGTCCTCCAGGACTGGACATATCTGCACTCTCTGCTGAC 972
|
|
|
QY 321 PheSerArgTyrPheGlnTyrGluGlySerLeuThrThrProProCysAlaGlnGlyVal 340
|
|
|
Db 973 TTCAAGCCGCTACTTCCATATGAGGGGTCTCTGACTACACCGCCCTGTGCCCCAGGGTGC 1032
|
|
|
QY 341 IleTyrThrValPheAsnGlnThrValMetLeuSerAlaLysGlnLeuHisThrLeuSer 360
|
|
|
Db 1033 ATCTGGACTGTGTTTAACCCAGACAGTGTGAGTGCTAAGCAGCTCCACACCCCTCTCT 1092
|
|
|
QY 361 AspThrLeuTyrGlyProGlyAspSerArgLeuGlnLeuAsnPheArgAlaThrGlnPro 380
|
|
|
Db 1093 GACACCTGTGGGACCTGTGACTCTCGGCTACAGCTGACTTCCGAGCGACGACGCT 1152
|
|
|
QY 381 LeuAsnGlyArgValIleGluAlaSerPheProAlaGlyValAspSerSerProArgAla 400
|
|
|
Db 1153 TTGAATGGGCGAGTGTAGGCGCTCTTCCCTGCTGGAGTGACAGCAGTCTCTCGGGCT 1212
|
|
|
QY 401 AlaGluProValGlnLeuAsnSerCysLeuAlaAlaGlyAspIleLeuAlaLeuValPhe 420
|
|
|
Db 1213 GCTGAGCCAGTCCAGCTGAATTCCTGCTGGCTGCTGGTGACATCTTAGCCCTGGTTTTT 1272
|
|
|
QY 421 GlyLeuLeuPheAlaValThrSerValAlaPheLeuValGlnMetArgArgGlnHisArg 440
|
|
|
Db 1273 GGCTCTCTTTTGTGCTACCGGCTCGGCTCTTGTGACATGAGAGGCGACGACAGA 1332
|
|
|
QY 441 ArgGlyThrIysGlyGlyValSerTyrArgProAlaGluValAlaGluThrGlyAla 459
|
|
|
Db 1333 AGGGGAACCAAGGGGTGTGAGCTACCGCCAGCAGAGTAGCCGAGACTGGAGCC 1389

RESULT 8
LOCUS AR171392 1522 bp DNA linear PAT 17-DEC-2001
DEFINITION Sequence 1 from patent US 6297041.
```

```
ACCESSION AR171392
VERSION AR171392.1 GI:17910342
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 1522)
AUTHORS Zavada,J., Pastorekova,S. and Pastorek,J.
TITLE MN gene and protein
JOURNAL Patent: US 6297051-A 1 02-OCT-2001;
FEATURES Location/Qualifiers
source 1..1522
/organism="unknown"
/mol_type="unassigned DNA"

ORIGIN
Alignment Scores:
Pred. No.: 6.2e-103 Length: 1522
Score: 2424.00 Matches: 459
Percent Similarity: 100.0% Conservatives: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 6 Gaps: 0

US-09-967-237B-2 (1-459) x AR171392 (1-1522)

QY 1 MetAlaProLeuCysProSerProTrpLeuProLeuLeuLeuProAlaProGly 20
DB 13 ATGGCTCCCTGTGCCAGCCCTGGCTCCCTCTGTGTGATCCGGCCCTGTCCAGGC 72
QY 21 LeuThrValGlnLeuLeuSerLeuLeuLeuMetProValHisProGlnArgLeu 40
DB 73 CTCACTGTGCAATGCTGCTGTCACTGCTGCTTCTGATGCTGTCCATCCCCAGAGGTTG 132
QY 41 ProArgMetGlnGluAspSerProLeuGlyGlySerSerGlyGluAspAspProLeu 60
DB 133 CCCCGGATGCAGAGGATTCCTCCCTTGGGAGGAGGCTCTTCTCGGAGAGATGATCCACTG 192
QY 61 GlyGluGluAspLeuProSerGluGluAspSerProArgGluGluAspProGlyGlu 80
DB 193 GCGCAGAGGAGTCTGCCAGTAGAGGATTCACCCAGAGAGGAGATCCACCCGAGAG 252
QY 81 GluAspLeuProGlyGluGluAspLeuProGlyGluGluAspLeuProGluValysPro 100
DB 253 GAGGATCTACTCGAGAGGAGGATCTACCTGGAGAGGAGGATCTACTGAAGTTAAGCCT 312
QY 101 LysSerGluGluGlySerLeuLysLeuGluAspLeuProThrValGluAlaProGly 120
DB 313 AATCAGAGAGAGGGCTCCCTGAGTTAGAGATCTACTACTGTGAGGCTCTCTGGA 372
QY 121 AspProGlnGluProGlnAsnAlaHisArgAspLysGluGlyAspAspGlnSerHis 140
DB 373 GATCCTCAAGAACCCAGATAATATGCCCAAGGACAAAGAGGGGATGACACAGAGTCAT 432
QY 141 TrpArgTyGlyGlyAspProProTrpProArgValSerProAlaCysAlaGlyArgPhe 160
DB 433 TGGCGCTATGAGCGCAGCCCTCCCTGGCCCCGGGTGTCCACAGCTGCGGGGGCCGCTTC 492
QY 161 GlnSerProValAspLysArgProGlnLeuAlaPheCysProAlaLeuArgProLeu 180
DB 493 CAGTCCCGGTGGATATCCGCCCCAGCTGCGCGCTTCTGCGGGCCCTGCGCCCCCTG 552
QY 181 GluLeuLeuGlyPheGlnLeuProProLeuProGluLeuArgLeuArgAsnAnglyHis 200
DB 553 GAATCTCTGGGCTTCCAGCTCCCGCGCTCCAGAACTGCGCTCCGCGCAACATGGCCAC 612
QY 201 SerValGlnLeuThrLeuProProGlyLeuGluMetAlaLeuGlyArgGluTyTr 220
DB 613 AGTGTGCAACTGACCTTGCCTCTCTGGGCTAGAGATGCTCTGGGTCCTGGGGCGGAGTAC 672
QY 221 ArgAlaLeuGlnLeuHisLeuHisTrpGlyValAlaGlyArgProGlySerGluHisThr 240
DB 673 CGGGCTCTGCAGCTGCATCTGCATCTGGGGGGCTGCAGGTCTCGGGCTCGGAGCACACT 732

241 ValGluGlyHisArgPheProAlaGluLeuHisValValHisLeuSerThrAlaPheAla 260
DB 733 GTGGAGGGCCACCGTTTCCCTGCCAGATCCACGTGGTTTCACCTCAGCACCCGCTTTGCC 792
QY 261 ArgValAspGluAlaLeuGlyArgProGlyGlyLeuAlaValLeuAlaAlaPheLeuGlu 280
DB 793 AGAGTTGACGAGCCCTTGGGGGCCCGCGGAGCCCTGGCCGTGTGGCCGCCCTTCTGGAG 852
QY 281 GluGlyProGluGluAsnSerAlaTyTrGluGlnLeuLeuLeuSerArgLeuGluLeuAla 300
DB 853 GAGGGCCCGGAAGAAAACAGTGCTTATGAGCAGTTGCTGTCTCGCTTGAAGAAATCGCT 912
QY 301 GluGluGlySerGluThrGlnValProGlyLeuAspIleSerAlaLeuLeuProSerAsp 320
DB 913 GAGGAAGGCTCAGAGACTCAGGTCCAGGACTGGACATATCTGCACTCTCGCTCCCTTGAC 972
QY 321 PheSerArgTyTrPheGlnTyTrGluGlySerLeuThrProProCysAlaGlnGlyVal 340
DB 973 TTCAGCCGCTACTTCCATATGAGGGGTCTCTGACTACACCGCCCTGTGCCAGGGTGTTC 1032
QY 341 IleTrpThrValPheAsnGlnThrValMetLeuSerAlaLysGlnLeuHisThrLeuSer 360
DB 1033 ATCTGGACTGTGTTTAAACCAGACAGTGATGCTGAGTGCTTAAGCAGCTCCACACCCCTCTCT 1092
QY 361 AspThrLeuTrpGlyProGlyAspSerArgLeuGlnLeuAsnPheArgAlaThrGlnPro 380
DB 1093 GACACCTGTGGGACCTGCTGACTCTCGGCTACAGCTGAACCTCCGAGCGCAGCGCCT 1152
QY 381 LeuAsnGlyArgValIleGluAlaSerPheProAlaGlyValAspSerSerProArgAla 400
DB 1153 TTGAATGGCGGAGTGATGAGGCTCTCTTCCCTGCTGGAGTGGACAGCAGTCTCTCGGGCT 1212
QY 401 AlaGluProValGlnLeuAsnSerCysLeuAlaAlaGlyAspIleLeuAlaLeuValPhe 420
DB 1213 GCTGAGCCAGTCCAGCTGAATTCCTGCTGGCTGCTGCTGGTGACATCTAGCCCTGGTTTTT 1272
QY 421 GlyLeuLeuPheAlaValThrSerValAlaPheLeuValGlnMetArgArgGlnHisArg 440
DB 1273 GGCCTCTCTTTTGTGTCCACAGCGTCGCGTTCCTTGTGCAGATGAGAAGGCGACAGAGA 1332
QY 441 ArgGlyThrLysGlyGlyValSerTyArgProAlaGluValAlaGluThrGlyAla 459
DB 1333 AGGGGACCAAAAGGGGTGTGAGTACC GCCCAGCAGAGGTAGCCGAGAGTGGAGCC 1389

RESULT 9
AR171563 1522 bp DNA linear PAT 17-DEC-2001
LOCUS Sequence 1 from patent US 6297051.
DEFINITION AR171563
ACCESSION AR171563
VERSION AR171563.1 GI:17910513
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 1522)
AUTHORS Zavada,J., Pastorekova,S. and Pastorek,J.
TITLE MN gene and protein
JOURNAL Patent: US 6297051-A 1 02-OCT-2001;
FEATURES Location/Qualifiers
source 1..1522
/organism="unknown"
/mol_type="unassigned DNA"

ORIGIN
Alignment Scores:
Pred. No.: 6.2e-103 Length: 1522
Score: 2424.00 Matches: 459
Percent Similarity: 100.0% Conservatives: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 6 Gaps: 0
```



|||||  
73 CTCACTGTGCAACTGCTGCTGTCACTGCTCTTCTGATGCTCTGTCAATCCCCAGAGGTTG 132  
QY  
41 ProArgMetGlnGluAspSerProLeuGlyGlySerSerGlyGluAspProLeu 60  
Db  
133 CCCGGATGCAGAGGATTCCCTTTGGGAGGAGGCTTCTTGGGGAAGATGACCCACTG 192  
QY  
61 GlyGluGluAspLeuProSerGluGluAspSerProArgGluGluAspProProGlyGlu 80  
Db  
193 GGGAGAGGAGGATCTGCCAGTGAAGAGGATTCCACCAGAGAGGAGGATCCACCCGGAGAG 252  
QY  
81 GluAspLeuProGlyGluGluAspLeuProGlyGluGluAspLeuProGluValLysPro 100  
Db  
253 GAGGATCTACTCGAGAGGAGGATCTACCTTGGAGAGGAGGATCTACTGGAAGTTAAGCCT 312  
QY  
101 LysSerGluGluGluGlySerLeuLysLeuGluAspLeuProThrValGluAlaProGly 120  
Db  
313 AAATCAGAGAGAGGCTCCCTGAAAGTTAGAGGATCTACTACTGTTGAGGCTCTCTGGA 372  
QY  
121 AspProGlnGluProGlnAenAenAlaHisArgAspLysGluGlyAspAspGlnSerHis 140  
Db  
373 GATCCTCAAGAACCCAGATAATAGCCCAAGGACAAAGAGAGGAGGATGACACAGAGTCAT 432  
QY  
141 TrpArgTyrGlyGlyAspProProTTPProArgValSerProAlaCyValGlyArgPhe 160  
Db  
433 TGGCGCTATGGAGGACACCGCCCTGCGCCCGGGGTGTCCCAAGCTGCGCGGGCCGCTTC 492  
QY  
161 GlnSerProValAspIleArgProGlnLeuAlaPheCysProAlaLeuArgProLeu 180  
Db  
493 CAGTCCCGGTGATATCCGCCCCCAGCTGCGCCCTTCTGCGCGGCGCTGCGCCCCCTG 552  
QY  
181 GluLeuLeuGlyPheGlnLeuProProLeuProLeuProGluLeuArgLeuArgAenGlyHis 200  
Db  
553 GAACTCCTGGGCTTCCAGCTCCCGCGCTCCAGAACTGCGCCTGCGCAACATGGCCAC 612  
QY  
201 SerValGlnLeuThrLeuProProGlyLeuGluMetAlaLeuGlyProGlyArgGluTyr 220  
Db  
613 AGTGTGAACGTGACCTTCCCTCTGCGCTAGAGATGGCTCTGGGTCCCGGCGGAGTAC 672  
QY  
221 ArgAlaLeuGlnLeuHisLeuHisTrpGlyAlaAlaGlyArgProGlySerGluHisThr 240  
Db  
673 CGGGCTCTGAGTGCATCTGCATCTGGGGGCTGCAGGTCTGTCGGGCTCGAGACACT 732  
QY  
241 ValGluGlyHisArgPheProAlaGluLysValHisValHisLeuSerThrAlaPheAla 260  
Db  
733 GTGGAAGGCCACCGCTTCCCTGCGAGATCCAGCTGGTTCCACTCAGCACCGCTTTGGC 792  
QY  
261 ArgValAspGluAlaLeuGlyArgProGlyGlyLeuAlaValLeuAlaAlaPheLeuGlu 280  
Db  
793 AGAGTTGACAGGCGCTTGGGGCGCCCGGAGGCTGGCCGTGTGGCGGCTTTCTGGAG 852  
QY  
281 GluGlyProGluGluAenSerAlaTyrGluGlnLeuLeuSerArgLeuGluLysAla 300  
Db  
853 GAGGGCCCGGAAGAAACAGTGCTTATGAGCAGTTGCTGCTCGCTTGGAAAGAAATCGCT 912  
QY  
301 GluGluGlySerGluThrGlnValProGlyLeuAspIleSerAlaLeuLeuProSerAsp 320  
Db  
913 GAGGAAGGCTCAGAGACTCAGGTCCAGGATCCGACATATCTGCATCTCTGCGCTCTGAC 972  
QY  
321 PheSerArgTyrPheGlnTyrGluGlySerLeuThrThrProProCysAlaGlnGlyVal 340  
Db  
973 TTCAGCGGCTACTTCCATATGAGGGGTCTCTGACTACACCGCCCTGTGCCAGGGGTGTC 1032  
QY  
341 IleTrpThrValPheAenGlnThrValMetLeuSerAlaLysGlnLeuHisThrLeuSer 360  
Db  
1033 ATCTGGACTGTGTTTAACAGACAGTGTGCTGAGTGTGAAGCAGCTCCACACCCCTCTCT 1092  
QY  
361 AspThrLeuTrpGlyProGlyAspSerArgLeuGlnLeuAenPheArgAlaThrGlnPro 380  
Db  
1093 GACACCCCTGTGGGACCTGTGTGACTCTCGGTACAGCTGAACCTTCGAGCGACGACGCT 1152  
QY  
381 LeuAenGlyArgValLysGluAlaSerPheProAlaGlyValAspSerProArgAla 400  
|||||

Db 1153 TTCAATGGCGGAGTGATTGAGGCCTCTTCCCTGCTGGAGTGACACAGCAGTCTCTCGGCT 1212  
QY 401 AlaGluProValGlnLeuAenSerCysLeuAlaAlaGlyAspIleLeuAlaLeuValPhe 420  
Db 1213 GCTGAGCAGTCCAGCTGAATTCTGCTGGCTGCTGGTGACATCTAGCCCTGGTTTTT 1272  
QY 421 GlyLeuLeuPheAlaValThrSerValAlaPheLeuValGlnMetArgArgGlnHisArg 440  
Db 1273 GGCCTCTTTTGTGCTGTCACAGCGTCCGCTTCTTGTGAGATGAGAAGGCGACAG 1332  
QY 441 ArgGlyThrLysGlyGlyValSerTyrArgProAlaGluValAlaGluThrGlyAla 459  
Db 1333 AGGGAGAACCAAGGGGTGTGAGCTACCGCCACGACAGAGGTAGCCGAGACTGGAGCC 1389  
RESULT 11  
AR569592  
LOCUS AR569592 1522 bp mRNA linear PAT 14-DEC-2004  
DEFINITION Sequence 1 from patent US 6770438.  
ACCESSION AR569592  
VERSION AR569592.1 GI:56570221  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.  
REFERENCE  
1 (bases 1 to 1522)  
AUTHORS Zavada,J., Pastorekova,S. and Pastorek,J.  
TITLE MN gene and protein  
JOURNAL Patent: US 6770438-A 1 03-AUG-2004;  
Institute of Virology, Slovak Academy of Sciences; Bratislava;  
CZX;  
FEATURES  
source  
1..1522  
/organism="unknown"  
/mol\_type="mRNA"  
ORIGIN  
Alignment Scores:  
Pred. No.: 6,2e-103 Length: 1522  
Score: 2424.00 Matches: 459  
Percent Similarity: 100.0% Conservative: 0  
Best Local Similarity: 100.0% Mismatches: 0  
Query Match: 100.0% Indels: 0  
DB: 6 Gaps: 0  
US-09-967-237B-2 (1-459) x AR569592 (1-1522)  
QY 1 MetalProLeuCysProSerProTrpLeuProLeuLeuLeuProAlaProGly 20  
Db 13 ATGGCTCCCTGTGCCCGCCAGCCCTGGCTCCCTCTGTTGATCCCGGCCCTGTCTCCAGGC 72  
QY 21 LeuThrValGlnLeuLeuSerLeuLeuLeuMetProValHisProGlnArgLeu 40  
Db 73 CTCACCTGTGCAACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 132  
QY 41 ProArgMetGlnGluAspSerProLeuGlyGlyGlySerSerGlyGluAspProLeu 60  
Db 133 CCCCCGATGCAGAGGATTCCTCCCTTGGGAGGAGGCTCTTCTGGGAAAGATGCCACTG 192  
QY 61 GlyGluGluAspLeuProSerGluGluAspSerProArgGluGluAspProProGlyGlu 80  
Db 193 GGGAGAGGATCTGCCAGTGAAGAGGATTCACCCAGAGAGGAGGATTCACCCCGGAGAG 252  
QY 81 GluAspLeuProGlyGluGluAspLeuProGlyGluGluAspLeuProGluValLysPro 100  
Db 253 GAGGATCTACTCTGGAGAGGAGGATCTACCTGAGAGGAGGATCTACTGGAAGTTAAGCCT 312  
QY 101 LysSerGluGluGluGlySerLeuLysLeuGluAspLeuProThrValGluAlaProGly 120  
Db 313 AAATCAGAGAGAGGCTCCCTGAAAGTTAGAGGATCTACTACTGTTGAGGCTCTCTGGA 372  
QY 121 AspProGlnGluProGlnAenAenAlaHisArgAspLysGluGlyAspAspGlnSerHis 140  
Db 373 GATCCTCAAGAACCCAGATAATAGCCCAAGGACAAAGAGAGGAGGATGACACAGAGTCAT 432



```
Db 733 GTGGAAAGCCACCGTTTCCCTGCCGAGATCCAGTGCTTACCTCAGACCCGCTTTGGC 792
Qy 261 ArgValAspGluAlaLeuGlyArgProGlyGlyLeuAlaValLeuAlaPheLeuGlu 280
Db 793 AGAGTTGACGAGCCCTTGGGCGCCGGGAGGCTTGGCGTGTGGCCGCTTTCTGGAG 852
Qy 281 GluGlyProGluGluAsnSerAlaTyrGluGlnLeuSerArgLeuGluLufLeAla 300
Db 853 GAGGGCCCGGAAGAAACAGTGCTATGACAGTTCGTCTCGCTTGGAAAGAAATCGCT 912
Qy 301 GluGluGlySerGlnThrGlnValProGlyLeuAspIleSerAlaLeuLeuProSerAsp 320
Db 913 GAGGAAGGCTCAGAGACTCAGGTCCAGGACTGGACATATCTGCATCTCTGCCCTCTGAC 972
Qy 321 PheSerArgTyrPheGlnTyrGluGlySerLeuThrThrProProCysAlaGlnGlyVal 340
Db 973 TTCAGCCGCTACTTCCAATATGAGGGTCTCTGACTACACCGCCCTGTGCCAGGGTGTG 1032
Qy 341 IleTyrThrValPheAsnGlnThrValMetLeuSerAlaIysGlnLeuHisThrLeuSer 360
Db 1033 ATCTGAGCTGTGTTTAACACGACAGTGTCTGAGTGCTAAGCAGCTCCACACCCCTCT 1092
Qy 361 AspThrLeuTyrGlyProGlyAspSerArgLeuGlnLeuAsnPheArgAlaThrGlnPro 380
Db 1093 GACACCTGTGGGACCTGTGTACTCTCGCTACAGTGAATCTCCAGGCGACGACCT 1152
Qy 381 LeuAsnGlyArgValIleGluAlaSerPheProAlaGlyValAspSerProArgAla 400
Db 1153 TTGAATGGCGAGTGATGAGGCTCTCTCCCTGCTGGAGTGACAGAGTCTCTGGGCT 1212
Qy 401 AlaGluProValGlnLeuAsnSerCysLeuAlaAlaGlyAspIleLeuAlaLeuValPhe 420
Db 1213 GCTGAGCCAGTCCAGCTGAATCTCGCTGGCTGCTGGTGACATCTCAGCCCTGGT 1272
Qy 421 GlyLeuLeuPheAlaValThrSerValAlaPheLeuValGlnMetArgArgGlnHisArg 440
Db 1273 GGCTCTCTTTTCTGTCACCGCTGCTGCTTCTTGTGCAGATGAGAGGCGACAG 1332
Qy 441 ArgGlyThrIysGlyGlyValSerTyrArgProAlaGluValAlaGluThrGlyAla 459
Db 1333 AGGGGAACCAAGGGGTGTGAGCTACCGCCAGCAGAGAGTAGCCGAGACTGGAGCC 1389

RESULT 13
CQ834035 1552 bp DNA linear PAT 29-JUL-2004
LOCUS Sequence 71 from Patent EP1439393.
DEFINITION CQ834035
ACCESSION CQ834035.1 GI:50833640
VERSION 1
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominoidea; Homo.
REFERENCE 1
AUTHORS Aetle, J.H., Boardman, L.A., Bugart, L.J., Burgess, C.C., Catino, T.J.,
Dwivedi, P., Huntress, M., Johnson, K.A., Lewis, M.E., Maimonis, P.J.,
Myerow, S.H., Brown-Shimer, S.L., Thiagalingam, A., Thibodeau, S.N. and
Molino, G.A.
TITLE Detection methods using TIMP 1 for colon cancer diagnosis
JOURNAL Patent: EP 1439393-A 71 21-JUL-2004;
Bayer Healthcare LLC (US); MAYO FOUNDATION FOR MEDICAL EDUCATION
AND RESEARCH (US)
FEATURES
source Location/Qualifiers
1..1552
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
ORIGIN
Alignment Scores: 6.32e-103 Length: 1552
Pred. No.:
```

```
Score: 2424.00 Matches: 459
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 6 Gaps: 0
US-09-967-237B-2 (1-459) x CQ834035 (1-1552)
Qy 1 MetAlaProLeuCyProSerProTyrProLeuLeuLeuProAlaProAlaProGly 20
Db 43 ATGGTCTCCCTTGGCCAGCCCTGGCTCCCTCTGTGTGATCCGCGCCCTGTCTCAGGC 102
Qy 21 LeuThrValGlnLeuLeuSerLeuLeuLeuMetProValHisProGlnArgLeu 40
Db 103 CTCACCTGTGCAACTGTCTGTCTGCTGCTTCTGATGCCCTGTCCATCCCGAGGTTG 162
Qy 41 ProArgMetGlnGluAspSerProLeuGlyGlySerSerGlyGluAspAspProLeu 60
Db 163 CCCCCGATGCAGAGGATTTCCCTTGGGAGGAGGCTCTTCTGGGGAAGATGACCCACTG 222
Qy 61 GlyGluGluAspLeuProSerGluGluAspSerProArgGluGluAspProProGlyGlu 80
Db 223 GCGCAGGAGGATCTGCCAGTGAAGAGGATTCACCCAGAGAGGAGATCCACCCGAGAG 282
Qy 81 GluAspLeuProGlyGluGluAspLeuProGlyGluGluAspLeuProGluValIysPro 100
Db 283 GAGGATCTACCTGGAGAGGAGATCTACCTGGAGAGGAGATCTACCTGAGTTAGCCT 342
Qy 101 LysSerGluGluGlySerLeuLysLeuGluAspLeuProThrValGluAlaProGly 120
Db 343 AAATCAGAGAAGAGGGCTCCCTGAAGTTAGAGGATCTACTACTTGTGAGGCTCTCTGA 402
Qy 121 AspProGlnGluProGlnAsnAsnAlaHisArgAspLysGlyGlyAspAspGlnSerHis 140
Db 403 GATCTCTAAGAACCCAGAGATTAATGCCACAGGGAACAAAGAGGGATGACACAGAGTCA 462
Qy 141 TrpArgTyrGlyGlyAspProProTyrProArgValSerProAlaCysAlaGlyArgPhe 160
Db 463 TGGCGCTATGAGAGCGACCCCGCTGGCCCGGGTGTCCCGAGCTTGGCGGGCGCTTC 522
Qy 161 GlnSerProValAspIleArgProGlnLeuAlaPheCysProAlaLeuArgProLeu 180
Db 523 CAGTCCCGGTGGATATCCGCCGCCAGCTCGCGGCTTCTGCGCGGCTTGTGGCGGCGGTG 582
Qy 181 GluLeuLeuGlyPheGlnLeuProProLeuProGluLeuArgLeuArgAsnGlyHis 200
Db 583 GAACTCTCGGGCTTCCAGCTCCCGCGCTCCAGAACTCGCGCTGCGCAACAAATGGCCAC 642
Qy 201 SerValGlnLeuThrLeuProProGlyLeuGluMetAlaLeuGlyProGlyArgGluTyr 220
Db 643 AGTGTGCACTGACCTGCTCTGGGCTAGAGATGGCTCTGGGTCCCGGGGGAGTAC 702
Qy 221 ArgAlaLeuGlnLeuHisLeuHisTrpGlyAlaAlaGlyArgProGlySerGluHisThr 240
Db 703 CGGGCTCTCGAGCTGCATCTGCACCTGGGGGGCTGCGAGTCTGCTCGCGGCTCGGAGCAC 762
Qy 241 ValGluGlyHisArgPheProAlaGluLleHisValValHisLeuSerThrAlaPheAla 260
Db 763 GTGGAGGCGCACCGTTCCTCCCGAGATCCACGTGGTTCACCTCAGCACCCGCTTTGGCC 822
Qy 261 ArgValAspGluAlaLeuGlyArgProGlyGlyLeuAlaValLeuAlaPheLeuGlu 280
Db 823 AGAGTTGACAGGCTTGGGGCGCCCGGAGGCTGGCGTGTGGCCGCTTTCTGGAG 882
Qy 281 GluGlyProGluGluAsnSerAlaTyrGluGlnLeuLeuSerArgLeuGluLufLeAla 300
Db 883 GAGGGCCCGGAAGAAACAGTGCCTATGACAGTTCGTCTGCTCGCTTGGAAAGAAATCGCT 942
Qy 301 GluGluGlySerGlnThrGlnValProGlyLeuAspIleSerAlaLeuLeuProSerAsp 320
Db 943 GAGGAAGGCTCAGAGACTCAGGTCCAGGACTGGAATATCTGCACTCTCTGCCCTCTGAC 1002
Qy 321 PheSerArgTyrPheGlnTyrGluGlySerLeuThrThrProProCysAlaGlnGlyVal 340
```

```
Db 1003 TTGAGCGCTACTTCCATATAGGGGTCTCTGACTACACGCGCTGTGCGCAGGGTCTC 1062
Qy 341 IletTprThrValPheAenGlnThrValMetLeuSerAlaLysGlnLeuHisThrLeuSer 360
Db 1063 ATCTGGACTGTGTTAAACACAGACAGTGTAGTGTAGTGTAGTGTAGTGTAGTGTAGTGT 1122
Qy 361 AspThrLeuTprGlyProGlyAspSerArgLeuGlnLeuAenPheArgAlaThrGlnPro 380
Db 1123 GACACCCCTGTGGGACCTGTGTGACTCTCGGTACAGTGAACCTTCGAGGCGACGACCT 1182
Qy 381 LeuAenGlyArgValIleGluAaSerPheProAlaGlyValAspSerSerProArgAla 400
Db 1183 TTGATGGGCGAGTGTAGTGTAGGCTCTCTCCCTGCTGGAGTGACAGCAGCTCTCGGCT 1242
Qy 401 AlaGluProValGlnLeuAenSerCysLeuAlaGlyAspIleLeuAlaLeuValPhe 420
Db 1243 GCTGAGCAGTCCAGCTGAATTCCTGCTGGCTGTGCTGTGCTGTGCTGTGCTGTGCTGT 1302
Qy 421 GlyLeuLeuPheAlaValThrSerValAlaPheLeuValGlnMetArgArgGlnHisArg 440
Db 1303 GGCTCTCTTTTGTGTCACAGCGTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCT 1362
Qy 441 ArgGlyThrLysGlyGlyValSerTyrArgProAlaGluValAlaGluThrGlyAla 459
Db 1363 AGGGGAACCAAGGGGGTGTGACTACCGCCAGCAGAGTAGCGGAGACTGGAGCC 1419

RESULT 14
LOCUS AX330007
DEFINITION Sequence 516 from Patent WO0194629.
ACCESSION AX330007
VERSION AX330007.1 GI:18102985
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
REFERENCE
1 Young, P. E., Augustus, M., Carter, K. C., Ebner, R., Endress, G.,
Horrihan, S., Soppet, D. R. and Weaver, Z.
Cancer gene determination and therapeutic screening using signature
gene sets
JOURNAL Patent: WO 0194629-A 516 13-DEC-2001;
Avalon Pharmaceuticals (US)
FEATURES
source
1. 1552
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
ORIGIN
Alignment Scores:
Pred. No.: 6.32e-103 Length: 1552
Score: 2424.00 Matches: 459
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 6 Gaps: 0

US-09-967-237B-2 (1-459) x AX330007 (1-1552)
Qy 1 MetAlaProLeuCysProSerProTprLeuProLeuLeuIleProAlaProGly 20
Db 43 ATGGCTCCCTGTGCCCCAGCCCTGCTCCCTCTGTGTATCCGCGCCCTGTCTCCAGGC 102
Qy 21 LeuThrValGlnLeuLeuSerLeuLeuLeuMetProValHisProGlnArgLeu 40
Db 103 CTCACTGTGCACTGCTGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 162
Qy 41 ProArgMetGlnGluAspSerProLeuGlyGlySerSerGlyGluAspProLeu 60
```

```
163 CCCCAGTGCAGGAGGATTCCCTTTGGGAGGAGGCTCTTCTGGGGAAGATGATCACCACCTG 222
61 GlyGluGluAspLeuProSerGluGluAspSerProArgGluGluAspProProGlyGlu 80
223 GCGAGGAGGATCTGCCAGTGAAGAGGATTCACCCAGAGAGGAGGATTCACCCGGAGAG 282
81 GluAspLeuProGlyGluGluAspLeuProGlyGluGluAspLeuProGluValLysPro 100
283 GAGGATCTACCTGGAGAGGAGGATCTACCTGGAGAGGAGGATCTACCTGAAGTTAGCCT 342
101 LysSerGluGluGluGlySerLeuLysLeuGluAspLeuProThrValGluAlaProGly 120
343 AAATCAGAAGAGAGGCTCCCTGAAGTTAGAGGATCTACCTACTGTGTAGGCTCTCTGGA 402
121 AspProGlnGluProGlnAenAenAlaHisArgAspLysGluGlyAspAspGlnSerHis 140
403 GATCTCCAGAAACCCAGAAATATGCCACAGGAGCAAGAAGAGGAGATGACAGAGTCAT 462
141 TprArgTyrGlyGlyAspProTprProArgValSerProAlaCysAlaGlyArgPhe 160
463 TGGCGCTATGGAGCGACCCGCTTGGCCCGGGTGTCCCCAGCCTGGCGCGCGCTTC 522
161 GlnSerProValAspIleArgProGlnLeuAlaAaPheCysProAlaLeuArgProLeu 180
523 CAGTCCCGCTGTGATATCCGCCCCAGCTGCGCGCTTCTGCGCGCGCTTCTGCGCGCTG 582
181 GluLeuLeuGlyPheGlnLeuProProLeuProGluLeuArgLeuArgAenGlnHis 200
583 GAATCTCTGGGCTTCCAGCTCCCGCGCTCCAGAACTGGCGCTGGCAACATATGGCCAC 642
201 SerValGlnLeuThrLeuProProGlyLeuGluMetAlaLeuGlyProGlyArgGluTyr 220
643 AGTGTGCAACTGACCTTGCCTCTGGCTAGAGATGGCTCTGGGTCCCGCGCGGAGTAC 702
221 ArgAlaLeuGlnLeuHisLeuHisTprGlyAlaAaGlyArgProGlySerGluHisThr 240
703 CGGGCTCTGAGCTGCTATCTGCACTGGGGGCTGTCAGGTCTGTCGGGCTCGGAGCACACT 762
241 ValGluGlyHisArgPheProAlaGluIleHisValValHisLeuSerThrAlaPheAla 260
763 GTGGAAGGCCACCGTTCCTGCGCGAGATCCAGTGTTCACCTCAGCACCGCTTTTGCC 822
261 ArgValAspGluAlaLeuGlyArgProGlyGlyLeuAlaValLeuAlaAlaPheLeuGlu 280
823 AGAGTTGACGAGGCTTGGGGCGCCGCGAGGCTTGGCGTGTGGCGCTTCTCTGAG 882
281 GluGlyProGluGluAenSerAlaTyrGluGlnLeuLeuSerArgLeuGluIleAla 300
883 GAGGGCCCGGAAGAAACAGTGTCTATGAGCAGTGTCTGTCTGCTTGGAAAGAAATCGCT 942
301 GluGluGlySerGluThrGlnValProGlyLeuAspIleSerAlaLeuLeuProSerAsp 320
943 GAGGAAGGCTCAGAGACTCAGGTCCCGAGGACTGGACATATCTGCATCTCTCCCTCTGAC 1002
321 PheSerArgTyrPheGlnTyrGluGlySerLeuThrThrProCysAlaGlnGlyVal 340
1003 TTCAGCGCTACTTCCAAATATAGGGGTCTCTGACTACACCGCTTGTGCCAGGGGTGTC 1062
341 IleTprThrValPheAenGlnThrValMetLeuSerAlaLysGlnLeuHisThrLeuSer 360
1063 ATCTGGACTGTGTTAAACACAGACAGTGTGCTATGAGCAGTGTCTGTCTGCTTCCCTCTCT 1122
361 AspThrLeuTprGlyProGlyAspSerArgLeuGlnLeuAenPheArgAlaThrGlnPro 380
1123 GACACCCCTGTGGGAGACTGTGTGACTCTCGGCTTAAGTGAACCTTCGAGGCGACGACCT 1182
381 LeuAenGlyArgValIleGluAaSerPheProAlaGlyValAspSerSerProArgAla 400
1183 TTGATGGGCGAGTGTAGTGTAGGCTCTCTCCCTGCTGGAGTGACAGCAGCTCTCGGCT 1242
401 AlaGluProValGlnLeuAenSerCysLeuAlaGlyAspIleLeuAlaLeuValPhe 420
1243 GCTGAGCAGTCCAGCTGAATTCCTGCTGGCTGTGCTGTGCTGTGCTGTGCTGTGCTGT 1302
```

QY 421 GlyLeuLeuPheAlaValThrSerValAlaPheLeuValGlnMetArgArgGlnHisArg 440  
Db 1303 GGCCCTCTTTTGTGTACACGGTCCGGCTTCTTGTGCAGATGAGAGGCAGACAGA 1362

QY 441 ArgGlyThrIysGlyValSerTyrArgProAlaGluValAlaGluThrGlyAla 459  
Db 1363 AGGGGAACCAAGGGGTGTAGCTACCGCCAGCAGAGGTGCCGAGACTGGAGCC 1419

RESULT 15  
AX332607 1552 bp DNA linear PAT 09-JAN-2002  
LOCUS AX332607  
DEFINITION Sequence 3116 from Patent WO0194629.  
ACCESSION AX332607  
VERSION AX332607.1 GI:18123241  
KEYWORDS  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Homnidae; Homo.

REFERENCE 1  
AUTHORS Young, P.E., Augustus, M., Carter, K.C., Ebner, R., Endress, G.,  
Horrikan, S., Soppet, D.R. and Weaver, Z.  
TITLE Cancer gene determination and therapeutic screening using signature  
gene sets  
JOURNAL Patent: WO 0194629-A 3116 13-DEC-2001;  
Avalon Pharmaceuticals (US)  
FEATURES  
source  
1..1552  
Location/Qualifiers  
/organism="Homo sapiens"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:9606"

ORIGIN  
Alignment Scores:  
Pred. No.: 6,32e-103 Length: 1552  
Score: 2424.00 Matches: 459  
Percent Similarity: 100.0% Conservative: 0  
Best Local Similarity: 100.0% Mismatches: 0  
Query Match: 100.0% Indels: 0  
DB: 6 Gaps: 0

US-09-967-237B-2 (1-459) x AX332607 (1-1552)

QY 1 MetAlaProLeuCysProSerProTrrPLeuProLeuLeuIleProAlaProGly 20  
Db 43 ATGGCTCCCTGTGCCCCAGCCCTGGCTCCCTCTGTGTATCCCGCCCTGTCTCCAGGC 102

QY 21 LeuThrValGlnLeuLeuSerLeuLeuLeuLeuMetProValHisProGlnArgLeu 40  
Db 103 CTCACGTGTGCMACTGTCTGTCTACCTGCTCTTCTGTATGCTCTCCATCCCGAGGTTG 162

QY 41 ProArgMetGlnGluAspSerProLeuGlyGlySerSerGlyGluAspAspProLeu 60  
Db 163 CCCCAGATGACAGAGGATTCCTTGGAGGAGGCTCTTCTGGGAAGATGATCCACTG 222

QY 61 GlyGluGluAspLeuProSerGluGluAspSerProArgGluGluAspProGlyGlu 80  
Db 223 GGCAGAGAGGATCTGCCAGTGAAGAGGATTCACCCAGAGAGGAGATCCACCCGAGAG 282

QY 81 GluAspLeuProGlyGluGluAspLeuProGlyGluGluAspLeuProGluValIysPro 100  
Db 283 GAGGATCTACTGGAGAGGAGGATCTACCTGGAGAGGAGGATCTACTGAAGTTAAGCCT 342

QY 101 LysSerGluGluGluGlySerLeuLysLeuGluAspLeuProThrValGluAlaProGly 120  
Db 343 AAATCAGAGAAGAGGGCTCCCTGAAGTTAGAGATCTACTTACTGTTGAGGGCTCTCGA 402

QY 121 AspProGlnGluProGlnAenAenAlaHisArgAspLysGlyGluGlyAspAspGlnSerHis 140  
Db 403 GATCCTCAAGAAACCCAGATAATGCCCAAGGGAACAAAGAGGGGATGACCAAGAGTCAT 462

QY 141 TrpArgTyrGlyGlyAspProProTrrPProArgValSerProAlaCysAlaGlyArgPhe 160  
Db 463 TGGCGTATGAGAGGACCCCGCTTGGCCCCGGGTGTCCCCAGCCTGCGCGGCGCCTTC 522

QY 161 GlnSerProValAspIleArgProGlnLeuAlaAlaPheCysProAlaLeuArgProLeu 180  
Db 523 CAGTCCCCGGTGGATATCCGCCCCAGCTCGCGGCTTCTGCCCCGCTTGGCCCCCTTG 582

QY 181 GluLeuLeuGlyPheGlnLeuProProLeuProGluLeuArgLeuArgAenGlyHis 200  
Db 583 GAACTCTCTGGGCTTCCAGCTCCCGCGCTCCAGAACTGCGCTGCGCAACAATGGCCAC 642

QY 201 SerValGlnLeuThrLeuProProGlyLeuGluMetAlaLeuGlyProGlyArgGluTyr 220  
Db 643 AGTGTGCAACTGACCTGCTCTTGGCTTAGAGATGGCTCTGGGTCCCCGGGCGGAGTAC 702

QY 221 ArgAlaLeuGlnLeuHisLeuHisTrpGlyAlaAlaGlyArgProGlySerGluHisThr 240  
Db 703 CCGGCTCTGCAGCTGCATCTGCACCTGGGGGGTTCAGGTCTGTCGGGCTCGGAGCACACT 762

QY 241 ValGluGlyHisArgPheProAlaGluIleHisValHisLeuSerThrAlaPheAla 260  
Db 763 GTGGAAGGCCACCGTTCCTGCGAGATCCACGTGTTTCACTCAGCACCGCTTTGGC 822

QY 261 ArgValAspGluAlaLeuGlyArgProGlyGlyLeuAlaValLeuAlaPheLeuGlu 280  
Db 823 AGAGTTGACAGGCGCTTGGGGCGCCCGGAGGCGCTGGCGTGTGGTCCCGCTTCTGGAG 882

QY 281 GluGlyProGluGluAsnSerAlaTyrGluGlnLeuLeuSerArgLeuGluIleAla 300  
Db 883 GAGGGCCCGAAGAAAACAGTGCCTATGAGCAGTGTCTCTCGCTTGAAGAAATCGCT 942

QY 301 GluGluGlySerGluThrGlnValProGlyLeuAspIleSerAlaLeuLeuProSerAsp 320  
Db 943 GAGGAAGGCTCAGAGACTCAGGTCCCGAGACTGGACATATCTGCACTCTCGCTCTGAC 1002

QY 321 PheSerArgTyrPheGlnTyrGlySerLeuThrProProCysAlaGlnGlyVal 340  
Db 1003 TTCAGCCGCTACTTCCATATGAGGGGTCTCTGACTACACCGCCCTGTGCCAGGGTGC 1062

QY 341 IleTrrPThrValPheAsnGlnThrValMetLeuSerAlaLysGlnLeuHisThrLeuSer 360  
Db 1063 ATCTGGACTGTGTAAACCAGACAGTGTGTGCTGAAGCAGCTCCACACCCCTCTCT 1122

QY 361 AspThrLeuTrrPglyProGlyAspSerArgLeuGlnLeuAsnPheArgAlaThrGlnPro 380  
Db 1123 GACACCTCTGGGGACCTGTGACTCTCGGCTACAGCTGAACCTTCGAGCGAGCGACCT 1182

QY 381 LeuAsnGlyArgValIleGluAlaSerPheProAlaGlyValAspSerSerProArgAla 400  
Db 1183 TTGAATGGCGAGGTGATTGAGGCTCTCTTCCCTGTGGAGTGGACAGCAGTCTCTCGGCT 1242

QY 401 AlaGluProValGlnLeuAsnSerCysLeuAlaAlaGlyAspIleLeuAlaLeuValPhe 420  
Db 1243 GCTGAGCCAGTCCAGCTGAATTCCTGCTGGCTGTCTGGTGACATCTAGCCCTGGTTTT 1302

QY 421 GlyLeuLeuPheAlaValThrSerValAlaPheLeuValGlnMetArgGlnHisArg 440  
Db 1303 GGCCTCTTTTGTGTCAACCGCTCGCGTCTCTTGTGCAGATGAGAGGCAGACAGA 1362

QY 441 ArgGlyThrLysGlyGlyValSerTyrArgProAlaGluValAlaGluThrGlyAla 459  
Db 1363 AGGGGAACCAAGGGGTGTAGCTACCGCCAGCAGAGGTAGCCGAGACTGGAGCC 1419

Search completed: February 17, 2006, 16:48:49  
Job time : 7268 secs

GenCore version 5.1.7  
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: February 17, 2006, 14:33:11 : Search time 841 Seconds

(without alignments)  
3637.448 Million cell updates/sec

Title: US-09-967-237b-2

Perfect score: 2424

Sequence: 1 MAPLCSPWMLPLIPAPAG.....RRGTGGVSYPAEVAETGA 459

Scoring table:

BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 4996997 seqs, 3332346308 residues

Total number of hits satisfying chosen parameters: 9993994

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh  
-Q=/abs/ABSSWEB\_spool/US09967237/runat\_15022006\_182241\_27953/app\_query.fasta\_1  
-DB=N\_Geneseq -QFWT=fastap -SUFFIX=ring -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0  
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45  
-DOCALIGN=200 -THR\_SCORE=pcr -THR\_MAX=100 -THR\_MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFWT=ptco -NORM=ext -HEADSIZE=500 -MINLEN=0 -MAXLEN=2000000000 -HOST=abs03h  
-USER=US09967237 @CGN\_1\_1727 @runat\_15022006\_182241\_27953 -NCPU=6 -ICPU=3  
-NO MMAP -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG -DEV\_TIMEOUT=120  
-WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7  
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : N\_Geneseq\_21.\*

1: Geneseq1980s.\*  
2: Geneseq1990s.\*  
3: Geneseq2000s.\*  
4: Geneseq2001as.\*  
5: Geneseq2001bs.\*  
6: Geneseq2002as.\*  
7: Geneseq2002bs.\*  
8: Geneseq2003as.\*  
9: Geneseq2003bs.\*  
10: Geneseq2003cs.\*  
11: Geneseq2003ds.\*  
12: Geneseq2004as.\*  
13: Geneseq2004bs.\*  
14: Geneseq2005s.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2424	100.0	1380	12 ADL70155	Adl70155 Human car
2	2424	100.0	1392	12 ADP54000	Adp54000 Human car
3	2424	100.0	1519	7 ADS73098	Ads73098 Human kid
4	2424	100.0	1519	7 ADW41952	Adw41952 cDNA elev

5	2424	100.0	1522	2 AAT09186	Aat09186 MuTu puta
6	2424	100.0	1522	3 AAA16540	Aaa16540 Human MN
7	2424	100.0	1522	3 AAS2459	Aas2459 Human MN
8	2424	100.0	1522	12 ADG31413	Adg31413 Human MN
9	2424	100.0	1522	12 ADK41803	Adk41803 Human MN
10	2424	100.0	1522	14 ADZ64594	Adz64594 Human MN
11	2424	100.0	1552	6 ABL64779	Ab164779 Lung canc
12	2424	100.0	1552	6 ABL62179	Ab162179 Colon ade
13	2424	100.0	1552	6 ABL65416	Ab165416 Lung canc
14	2424	100.0	1552	6 ABL68346	Ab168346 Kidney ca
15	2424	100.0	1552	8 ABX76385	Abx76385 Lung canc
16	2424	100.0	1552	8 ABX76124	Abx76124 Lung canc
17	2424	100.0	1552	10 ADG89343	Adg89343 Cancer de
18	2424	100.0	1552	10 ACC72730	Acc72730 Human can
19	2424	100.0	1552	10 ABZ77284	Abz77284 Nucleotid
20	2424	100.0	1552	11 ADN39704	Adn39704 Cancer/an
21	2424	100.0	1552	11 ADN38987	Adn38987 Cancer/an
22	2424	100.0	1552	12 ADG31472	Adg31472 Human Car
23	2424	100.0	1552	12 ADQ17643	Adq17643 Human sof
24	2424	100.0	1552	12 ADQ29645	Adq29645 Human col
25	2424	100.0	1552	13 ADR24713	Adr24713 Breast ca
26	2424	100.0	1552	14 ADY61778	Ady61778 Human gen
27	2424	100.0	1658	12 ADQ22315	Adq22315 Human sof
28	2424	100.0	1552	13 ADQ87399	Adq87399 Human tum
29	2419	99.8	1833	4 AAH26551	Aah26551 DNA encod
30	2416	99.7	1572	14 AEA00062	Aea00062 Human TAT
31	2416	99.7	1572	14 AEA00582	Aea00582 Human TAT
32	2036	84.0	1397	2 AAQ48456	Aaq48456 MN CDNA c
33	1844.5	76.1	1089	12 ADL70159	Adl70159 Human car
34	1600	66.0	978	12 ADL70161	Adl70161 Glu-tagge
35	1596	65.8	1965	12 ADG31474	Adg31474 Murine ca
36	1576	65.0	978	12 ADL70163	Adl70163 Glu-tagge
37	1176.5	48.5	10897	2 AAT09187	Aat09187 MuTu puta
38	1176.5	48.5	10898	3 AAA16543	Aaa16543 Human MN
39	1176.5	48.5	10898	3 AAS2462	Aas2462 Human MN
40	1176.5	48.5	10898	12 ADG31415	Adg31415 Human MN
41	1176.5	48.5	10898	12 ADK41805	Adk41805 Human MN
42	1176.5	48.5	10898	12 ADP54001	Adp54001 Human car
43	1173.5	48.4	10898	14 ADZ64596	Adz64596 Human MN
44	1159	47.8	660	14 ACL57169	ACL57169 Human col
45	988	40.8	586	12 ADG31473	Adg31473 Human MN

#### ALIGNMENTS

RESULT 1  
ADL70155  
ID ADL70155 standard; cDNA; 1380 BP.  
XX  
AC ADL70155;  
XX  
DT 20-MAY-2004 (first entry)  
XX  
DE Human carbonic anhydrase isozyme CA IX coding sequence.  
XX  
KW Human; carbonic anhydrase; tumour; antigen; cytosstatic; gene; aa.  
XX  
OS Homo sapiens.  
XX  
FH Key  
FT CDS  
FT CDS  
FT /tag= b  
FT /product= "Carbonic anhydrase CA IX"  
FT sig\_peptide  
FT 1..110  
FT /tag= a  
FT mat\_peptide  
FT 111..1377  
FT /tag= c  
XX  
DN WO2004017923-A2.  
XX  
PD 04-MAR-2004.  
XX  
PF 25-AUG-2003; 2003WO-US026612.

XX	23-AUG-2002; 2002US-0405577P.	QY	121	AspProGlnGluProGlnAsnAsnAlaHisArgAspGlyGluGlyAspAspGlnSerHis	140
PR	22-AUG-2003; 2003US-00646934.	DB	361	GATCTCTCAAGAACCCCGAGTAATAATGCCACAGGGACAAAGAGGGGATGACACAGTCAAT	420
XX	(CHIR ) CHIRON CORP.	QY	141	TpArgTyrGlyGlyAspProProTtpProArgValSerProAlaCysAlaGlyArgPhe	160
XX	Wong J, Winter J, Lalehzadeh G, Warne R;	DB	421	TGGCGCTATGGAGGGGACCCCGCTTGGCCCGGGGTGCCAGCCTTGGCGGGCCGCTTC	480
XX	WPI; 2004-226730/21.	QY	161	GlnSerProValAspIleArgProGlnLeuAlaAlaPheCysProAlaLeuArgProLeu	180
XX	P-PSDB; ADL70156.	DB	481	CAGTCCCGGTGGATATCCGCCCGCAGCTCGCGCTTCTGCGCGCCCTTGGCGCCCTG	540
PT	Identifying an agent capable of inhibiting carbonic anhydrase (CA)	QY	181	GluLeuLeuGlyPheGlnLeuProProLeuProGluLeuArgLeuArgAsnGlyHis	200
PT	test agent with a cell expressing CA IX and assessing CA inhibitory	DB	541	GAACCTCTGGGCTTCCAGCTCCCGCGCTCCAGAACTCGCGCTGCGCAACAATGGCCAC	600
PT	activity of the test agent.	QY	201	SerValGlnLeuThrLeuProProGlyLeuGluMetAlaLeuGlyProGlyArgGluTyr	220
PS	Example 6; SEQ ID NO 1; 82pp; English.	DB	601	AGTGTGCAACTGACCTGCTCTGGGCTAGAGATGGCTCTGGGTCCCGGGCGGAGTAC	660
CC	The present sequence is the coding sequence for human carbonic anhydrase	QY	221	ArgAlaLeuGlnLeuHisLeuHisIleArgGlyAlaAlaGlyArgProGlySerGluHisThr	240
CC	isozyme IX (CA IX), an N-glycosylated membrane-associated glycoprotein	DB	661	CGGGCTCTGCAGCTGCATCTGCACCTGGGGGGCTGCAGGTGCTCGGGCTCGGAGCACACT	720
CC	also known as the MN or G250 tumour-associated antigen. The invention	QY	241	ValGluGlyHisArgPheProAlaGluIleHisValValHisLeuSerThrAlaPheAla	260
CC	relates to compositions and methods useful in inhibiting CA IX+	DB	721	GTGGAGGCGCCCGCTTCCCTCCGAGATCCAGTGGTTCACCTCAGCACCGGCTTTGCC	780
CC	preneoplastic or neoplastic cells. The inhibitors are especially	QY	261	ArgValAspGluAlaLeuGlyArgProGlyGlyLeuAlaValLeuAlaAlaPheLeuGlu	280
CC	antagonistic anti-CA IX antibodies and other inhibitory agents that	DB	781	AGATTGACGAGGCTTGGGGGCGCCCGGAGGCTGGCCGTGTGGCCGCTTCTGGAG	840
CC	target the carbonic anhydrase activity of CA IX on these cells. The	QY	281	GluGlyProGluGluAsnSerAlaTyrGluGlnLeuLeuSerArgLeuGluIleAla	300
CC	antibodies, or their antigen-binding fragments, are specifically reactive	DB	841	GAGGGCCCGAGAGAAACAGTGCCTATGACAGTTCGTCTCTCGTTGGAAGAAATCGCT	900
CC	with an inhibitory epitope of CA IX. Screening assays for identifying	QY	301	GluGluGlySerGluThrGlnValProGlyLeuAspIleSerAlaLeuLeuProSerAsp	320
CC	such inhibitory agents are provided. The antibodies, their antigen-	DB	901	GAGGAAGGCTCAGAGACTCAGGTCCAGGACTGGACATATCTGCATCTCGCCCTCTGAC	960
CC	binding fragments, and other inhibitory agents are useful in the	QY	321	PheSerArgTyrPheGlnTyrGluGlySerLeuThrThrProProCysAlaGlnGlyVal	340
CC	treatment of cancers characterized by the expression of CA IX, such as:	DB	961	TTCAGCCGCTACTTCCCAATATGAGGGGTCTCTGACTACACCGCCCTGTGCCAGGGTGTC	1020
CC	head and neck cancers; gynaecological cancers including ovarian,	QY	341	IleTtpThrValPheAsnGlnThrValMetLeuSerAlaGlyGlnLeuHisThrLeuSer	360
CC	cervical, vaginal, endometrial and vulval cancers as well as	DB	1021	ATCTGGACTGTGTTTAAACAGACAGTGTGCTGAGTGTCTAAGCAGCTCCACACCTCTCT	1080
CC	gynaecological precancerous conditions such as metaplastic cervical	QY	361	AspThrLeuTtpGlyProGlyAspSerArgLeuGlnLeuAsnPheArgAlaThrGlnPro	380
CC	tissues and condylomas; gastrointestinal cancers such as stomach, colon	DB	1081	GACACCTGTGGGACCTGTGTGACTCTCGGCTACAGTGAACATTCGAGCGAGCGAGCTT	1140
CC	and oesophageal cancers; urinary tract cancer such as bladder and kidney	QY	381	LeuAsnGlyArgValIleGluAlaSerPheProAlaGlyValAspSerSerProArgAla	400
CC	cancers; skin cancer; liver cancer; prostate cancer; lung cancer; and	DB	1141	TTGAATGGCGGAGTGATTGAGGCTCTTCCCTCGCTGGAGTGGACAGCAGTCTCTCGGCT	1200
CC	breast cancer.	QY	401	AlaGluProValGlnLeuAsnSerCysLeuAlaAlaGlyAspIleLeuAlaLeuValPhe	420
XX	Sequence 1380 BP; 257 A; 428 C; 420 G; 275 T; 0 U; 0 Other;	DB	1201	GCTGAGCAGTCCAGCTGAATCTCTGCTGGGTGCTGGTGACATCTAGCCCTGGTTTTT	1260
Alignment Scores:		QY	421	GlyLeuLeuPheAlaValThrSerValAlaPheLeuValGlnMetArgArgGlnHisArg	440
Pred. No.:	1,67e-130	DB	1261	GGCTCTTTTGTGCTGACAGCGTGGGTCTCTTGTGAGATGAGAGGCGAGCAGA	1320
Score:	2424.00	QY	441	ArgGlyThrLysGlyGlyValSerTyrArgProAlaGluValAlaGluThrGlyAla	459
Percent Similarity:	100.0%	DB	1321	AGGGGACCAAGGGGGTGTGAGCTACCGCCAGCAGAGAGTAGCCGAGACTGGAGCC	1377
Best Local Similarity:	100.0%	RESULT 2			
Query Match:	100.0%	ID	ADP54000	standard; DNA; 1392 BP.	
DB:	12	XX	ADP54000		
US-09-967-237B-2 (1-459) x ADL70155 (1-1380)					
QY	1				
DB	1				
QY	21				
DB	61				
QY	41				
DB	121				
QY	61				
DB	181				
QY	81				
DB	241				
QY	101				
DB	301				





```

Db 610 AGTGTGCACTGACCTGCTCTGCTGGCTAGAGATGCTCTGGTCCCGGGGAGTAC 669
Qy 221 ArgAlaLeuGlnLeuHisHsfPrpGlyAlaAlaGlyArgProGlySerGluHisThr 240
Db 670 CGGGCTCTGCAGCTGATCTGCACATCGGGGGCTGCAGGTCTGTCGGGCTCGGAGCACAT 729
Qy 241 ValGluGlyHisArgPheProAlaGluIleHisValValHisLeuSerThrAlaPheAla 260
Db 730 GTGGAAGGCCACCGTTTCTCCGCGAGATCCAGTGTTCACCTCAGCACCGCTTTGGCC 789
Qy 261 ArgValAspGluAlaLeuGlyArgProGlyGlyLeuAlaValLeuAlaPheLeuGlu 280
Db 790 AGAGTTGACGAGGACCTTGGGGCCCGGGAGGCTGCGCGTGTGGCCGCTTTCTGGAG 849
Qy 281 GluGlyProGluGlnAsnSerAlaTyrGluGlnLeuSerArgLeuGluIleAla 300
Db 850 GAGGGCCCGGAGAAACAGTGCCTATGAGCAGTTGCTGTCTCGCTTGAAGAATAATCGCT 909
Qy 301 GluGlnGlySerGluThrGlnValProGlyLeuAspIleSerAlaLeuLeuProSerAsp 320
Db 910 GAGGAAGGCTCAGAGACTCAGGTCCAGGACTGGACATATCTGCATCTCTGCCCTCTGAC 969
Qy 321 PheSerArgTyrPheGlnTyrGluGlySerLeuThrThrProProCysAlaGlnGlyVal 340
Db 970 TTCAGCGCTACTTCCATATGAGGGGCTCTCTGACTACACCGCCCTGTGCCAGGGTGTCT 1029
Qy 341 IleTrpThrValPheAsnGlnThrValMetLeuSerAlaLysGlnLeuHisThrLeuSer 360
Db 1030 ATCTGCACTGTGTTAAACACAGACAGATGATGCTGAGTGTAAAGCAGCTCCACACCTCTCT 1089
Qy 361 AspThrLeuTrpGlyProGlyAspSerArgLeuGlnLeuAsnPheArgAlaThrGlnPro 380
Db 1090 GACACCTGTGGGACCTGTGTACTCTCGCTCAGCTGAACATTCGAGGAGCGACGCTCT 1149
Qy 381 LeuAsnGlyArgValIleGluAlaSerPheProAlaGlyValAspSerProArgAla 400
Db 1150 TTGAATGGGCGAGTGATGAGGCTCTTCCCTGCTGGAGTGACAGAGCTCTCGGGCT 1209
Qy 401 AlaGluProValGlnLeuAsnSerCysLeuAlaAlaGlyAspIleLeuAlaLeuValPhe 420
Db 1210 GCTGAGCCAGTCCAGCTGAATCTCGCTGGCTGCTGGTGACATCTCCTAGCCCTGTGTTTT 1269
Qy 421 GlyLeuLeuPheAlaValThrSerValAlaPheLeuValGlnMetArgArgGlnHisArg 440
Db 1270 GGGCTCTCTTTTGTGTACACAGCGTCCGCTTCTTGTGAGATGAGAGGACGACAGAGA 1329
Qy 441 ArgGlyThrLysGlyValSerTyrArgProAlaGluValAlaGluThrGlyVala 459
Db 1330 AGGGGAACCAAGGGGTGTGAGCTACCCCGCAGAGAGTAGCCGAGACTGGAGCC 1386

RESULT 4
ID ADW41952
XX ADW41952 standard; cDNA; 1519 BP.
XX AC
XX ADW41952;
XX
XX 24-MAR-2005 (first entry)
XX
XX cDNA elevated in kidney tumor cells SEQ ID NO 1695.
XX
XX gene therapy; diagnosis; pharmaceutical; cancer; cytostatic;
XX genitourinary disease; neoplasm; renal tumor; gene; ss.
XX
XX Unidentified.
XX
XX W0200274237-A2.
XX
XX 26-SEP-2002.
XX
XX 19-MAR-2002; 2002WO-US010055.
XX
XX 19-MAR-2001; 2001US-0277245P.
XX
PR

```

```

PR 21-DEC-2001; 2001US-0343340P.
XX
XX (CORI-) CORIXA CORP.
XX
XX Algate PA, Mannion J, Gaiger A, Gordon B, Harlocker SL;
XX
XX WPI; 2002-759855/82.
XX
XX New isolated polynucleotides and polypeptides, useful for detecting the
XX presence of, and treating cancer, particularly kidney cancer by
XX stimulating T-cells specific for a tumor protein, and stimulating immune
XX response in a patient.
XX
XX Claim 13; SEQ ID NO 1695; 252bp; English.
XX
XX The invention describes a new isolated polynucleotide (I) comprising: any
XX one of the 1855 sequences identified in the specification; complements or
XX degenerate variants of (a); sequences consisting of at least 20
XX contiguous residues of (a); sequences that hybridize to (a) under highly
XX stringent conditions; or sequences having at least 75 or 90% sequence
XX identity to (a). Also described are: detecting (M1) or determining the
XX presence of cancer in a patient; stimulating (M2) an immune response in a
XX patient; treating (M3) kidney cancer in a patient; an isolated
XX polypeptide (II) encoded by (I) and comprising, or having at least 70 or
XX 90% sequence identity to, any one of the 8 sequences identified in the
XX specification; an expression vector (III) comprising (I) operably linked
XX to an expression control sequence; a host cell (IV) transformed or
XX transfected with (III); an isolated antibody (V) or its antigen-binding
XX fragment that specifically binds to (II); a fusion protein (VI)
XX comprising at least one (II); an oligonucleotide (VII) that hybridizes to
XX the nucleotide sequences cited above under highly stringent conditions;
XX stimulating (M4) and/or expanding T-cells specific for a tumor protein;
XX an isolated T-cell population (VIII) comprising the T-cells in (M4); a
XX composition (IX) comprising a first component such as a carrier or
XX immunostimulant and a second component comprising (I), the polypeptide
XX encoded by (I), an antibody or its antigen-binding fragment that
XX specifically binds to (II), (VI), or an antigen-presenting cell that
XX expresses the polynucleotide; and a diagnostic kit (X) comprising at
XX least one of the polynucleotide, or at least one antibody and a
XX detection reagent comprising a reporter group. The polynucleotides,
XX polypeptides, antibodies and antigen-presenting cells are useful for
XX detecting the presence of, and treating cancer, particularly kidney
XX cancer by stimulating and/or expanding T-cells specific for a tumor
XX protein, and stimulating immune response in a patient. This sequence
XX represents a kidney tumour cDNA, expression of which is increased in
XX kidney tumors.
XX
XX Sequence 1519 BP; 296 A; 456 C; 453 G; 314 T; 0 U; 0 Other;
XX
XX Alignment Scores:
XX Pred. No.: 1,84e-130 Length: 1519
XX Score: 2424.00 Matches: 459
XX Percent Similarity: 100.0% Conservative: 0
XX Best Local Similarity: 100.0% Mismatches: 0
XX Query Match: 100.0% Indels: 0
XX DB: 7 Gaps: 0
XX
XX US-09-967-237B-2 (1-459) x ADW41952 (1-1519)
XX
XX Qy 1 MetAlaProLeuCysProSerProTrpLeuProLeuLeuLeuProAlaProGly 20
XX
XX Db 10 ATGGCTCCCTGTGCGCCAGCCCTGGCTCCCTCTGTGATCCCGGCCCTGTCTCCAGGC 69
XX
XX Qy 21 LeuThrValGlnLeuLeuSerLeuLeuLeuMetProValHisProGlnArgLeu 40
XX
XX Db 70 CTCACGTGTCACTGCTGTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 129
XX
XX Qy 41 ProArgMetGlnGluAspSerProLeuGlyGlySerGlyGlyGluAspProLeu 60
XX
XX Db 130 CCCCCGATGACGAGGATTCCTCCCTTGGGGGAGGCTCTTCTGGGGAAGATGATCCCACTG 189
XX
XX Qy 61 GlyGluGluAspLeuProSerGluGluAspSerProArgGluGluAspProGlyGlu 80
XX

```

Db 190 GCGAGGAGGATCTGCCAGTGAAGAGGATTCTCCACGAGAGGAGGATCCACCCGAGAG 249  
Qy 81 GluAspLeuProGlyGluGluAspLeuProGlyGluGluAspLeuProGluValLysPro 100  
Db 250 GAGGATCTACCTGAGGAGGAGGATCTACCTGGAGAGGAGGATCTACTGAAGTTAGCCT 309  
Qy 101 LysSerGluGluGlySerLeuLysLeuGluAspLeuProThrValGluAlaProGly 120  
Db 310 AAATCAGAGAGAGGGCTCCCTCAAGTTAGAGGATCTACTACTGTTGAGGCTCTCTGGA 369  
Qy 121 AspProGlnLeuProGlnLeuAsnAlaHisArgAspLysGluGlyAspAspGlnSerHis 140  
Db 370 GATCCTCAAGAACCCCAAGATAATGCCACAGGACAAAGAGGGGATGACCAAGTGCAT 429  
Qy 141 TrpArgTyrGlyGlyAspProProTyrProArgValSerProAlaCysAlaGlyArgPhe 160  
Db 430 TGGCGCTATGAGGCGACCCGCTTGGCCCGGGTGTCCCCACCTCGCGGGCCGCTTC 489  
Qy 161 GlnSerProValAspLeuArgProGlnLeuAlaAlaPheCysProAlaLeuArgProLeu 180  
Db 490 CAGTCCCGGTGGATATCCGCCGCCAGCTCGCGCCCTTCTGCGCGGCTTGGCGCCCTG 549  
Qy 181 GluLeuLeuGlyPheGlnLeuProLeuProGluLeuArgLeuArgLeuAsnGlyHis 200  
Db 550 GAATCTCTGGGCTTCCAGCTCCCGCGCTCCCAAGAACTGGCGCTCGCAACAATGGCCAC 609  
Qy 201 SerValGlnLeuThrLeuProProGlyLeuGluMetAlaLeuGlyProGlyArgGluTyr 220  
Db 610 AGTGTGCAACTGACCTTGCCTCTGGCTAGAGATGGCTCTGGGTCCCGGGGGAGTAC 669  
Qy 221 ArgAlaLeuGlnLeuHisLeuHisTrpGlyAlaAlaGlyArgProGlySerGluHisThr 240  
Db 670 CGGGCTCTGCAGCTGCATCTGCACATGGGGGCTGCAGGTCTCGGGCTCGGAGCACACT 729  
Qy 241 ValGluGlyHisArgPheProAlaGluIleHisValValHisLeuSerThrAlaPheAla 260  
Db 730 GTGGAAGGCCACCGTTTCCCTGCGGAGATCCAGTGGTTCCACTCAGCACCGCTTGTGC 789  
Qy 261 ArgValAspGluAlaLeuGlyArgProGlyGlyLeuAlaValLeuAlaAlaPheLeuGlu 280  
Db 790 AGAGTTGACGAGGCCCTTGGGGCGCCCGGAGGCTTGGCGGTGTGGCCGCTTCTGGAG 849  
Qy 281 GluGlyProGluLeuAsnSerAlaTyrGluGlnLeuSerArgLeuGluIleAla 300  
Db 850 GAGGGCCCGAAGAAACAGTGCCTATGACAGCTTGTCTGCTTGGAAAGAAATCGCT 909  
Qy 301 GluGluGlySerGluThrGlnValProGlyLeuAspLysSerAlaLeuLeuProSerAsp 320  
Db 910 GAGGAAGGCTCAGAGACTCAGGTCCAGGACTGGACATATCTGCATCTCTGCCCTCTGAC 969  
Qy 321 PheSerArgTyrPheGlnTyrGluGlySerLeuThrThrProCysAlaGlnGlyVal 340  
Db 970 TTCAGCGCTACTTCCAATATGAGGGTCTCTGACTACACCGCCCTGTGCCAGGGTGT 1029  
Qy 341 IleTrpThrValPheAsnGlnThrValMetLeuSerAlaLysGlnLeuHisThrLeuSer 360  
Db 1030 ATCTGAGCTGTGTAAACAGACAGTGTGCTGAGTGCTAAGCAGCTCCACACCTCTCT 1089  
Qy 361 AspThrLeuTrpGlyProGlyAspSerArgLeuGlnLeuAsnPheArgAlaThrGlnPro 380  
Db 1090 GACACCTGTGGGAGCTGTGTGACTCTGGCTACAGTGNACTTCCGAGGCGACGACCT 1149  
Qy 381 LeuAsnGlyArgValIleGluAlaSerPheProAlaGlyValAspSerSerProArgAla 400  
Db 1150 TTGAATGGCGAGTGTGAGTGGCTCTTCCCTGCTGGAGTGACAGCAGTCTCTCGGGCT 1209  
Qy 401 AlaGluProValGlnLeuAsnSerCysLeuAlaAlaGlyAspIleLeuAlaLeuValPhe 420  
Db 1210 GCTGAGCCAGTCCAGCTGAATCTCTCGCTGGCTGTGGTGACATCTAGCCCTGGTTTTT 1269  
Qy 421 GlyLeuLeuPheAlaValThrSerValAlaPheLeuValGlnMetArgArgGlnHisArg 440  
Db 1270 GGCCTCTTTTGTCTCACCAGCGTGGCTTCTTGTGCAGATGAGAGGCGACACAGA 1329

Qy 441 ArgGlyThrLysGlyGlyValSerTyrArgProAlaGluValAlaGluThrGlyAla 459  
Db 1330 AGGGGAACCAAGGGGTGTGAGCTACCGCCAGCAGAGGTAGCCGAGACTGGAGCC 1386

## RESULT 5

AAT09186  
ID AAT09186 standard; cDNA; 1522 BP.

XX AAT09186;

XX 25-MAR-2003 (revised)

DT 19-AUG-1996 (first entry)

XX MuTu putative oncogene MN cDNA.

DE MuTu; endogenous; cellular component; MN; HeLa cell; diagnosis;  
XX lymphocytic choriomeningitis virus; LCMV; putative oncogene; treatment;  
XX neoplastic; pre-neoplastic; disease; immunisation; ss.

OS Homo sapiens.

XX Key Location/Qualifiers  
PH 13..1392  
FT /\*tag= a

FT W09534650-A2.

PN 21-DEC-1995.

XX 15-JUN-1995; 95WO-US007628.

XX 15-JUN-1994; 94US-00260190.

PR 07-JUN-1995; 95US-00477504.

PR 07-JUN-1995; 95US-00481658.

PR 07-JUN-1995; 95US-00485049.

PR 07-JUN-1995; 95US-00485862.

PR 07-JUN-1995; 95US-00485863.

PR 07-JUN-1995; 95US-00486756.

PR 07-JUN-1995; 95US-00487077.

XX (CIBA ) CIBA CORNING DIAGNOSTICS CORP.

PA (VIRO-) INST VIROLOGY.

XX Zavada J, Pastorekova S, Pastorek J;

DR WPI; 1996-049679/05.

DR P-PSDB; AAR88058.

XX MN gene, protein and nucleic acid fragments - used as primers and probes  
PT in the detection of MN antigens and antibodies, and in the treatment of  
PT (pre)neoplastic disease.

XX Claim 1; Fig 1; 102pp; English.

XX The present sequence is the full length MuTu endogenous cellular  
CC component, MN, cDNA clone, which was isolated from lymphocytic  
CC choriomeningitis virus (LCMV) infected HeLa cells. Persistent LCMV, the  
CC exogenous MuTu transmissible agent (MX), infection increases the  
CC expression level of the MN gene. MN is a putative oncogene, and can  
CC therefore be used in the development of prods. for the diagnosis and  
CC treatment of neoplastic (NP), or pre-NP diseases. NP diseases can be  
CC treated using DNA antisense to MN transcribed mRNA, anti-MN protein  
CC antibodies can be used for the diagnosis NP or pre-NP diseases and a  
CC vaccine contg. immunogenic amounts of the MN protein can be used to  
CC immunise a vertebrate against a NP disease associated with MN antigen  
CC expression. (Updated on 25-MAR-2003 to correct PR field.)

XX Sequence 1522 BP; 297 A; 458 C; 453 G; 314 T; 0 U; 0 Other;

Alignment Scores: 1.85e-130 Length: 1522  
Pred. No.:

[illegible]





Db	1213	GCTGAGCAGTCCACAGCTGAATTCCTGCTGCCTGGTGACATCCCTAGCCCTGGTTTTT	1272
Qy	421	GlyLeuLeuPheAlaValThrSerValAlaPheLeuValGlnMetArgArgGlnHisArg	440
Db	1273	GGCTCTCTTTTGTGTGCCACAGAGTGGCTTCTTGTGCAGATGAGNAGGCACACAGA	1332
Qy	441	ArgGlyThrIysGlyGlyValSerTyrArgProAlaGluValAlaGluThrGlyAla	459
Db	1333	AGGGGAACCAAAAGGGGTGTGAGCTACCGCCACACAGAGGTAGCCGAGACTGGAGCC	1389
 RESULT 8 ADG31413			
ID	ADG31413	standard; cDNA; 1522 BP.	
XX	AC	ADG31413;	
XX	DT	26-FEB-2004 (first entry)	
XX	DE	Human MN cDNA SeqID 1.	
KW	KW	human; gene; ss; monoclonal antibody; MN/CA IX; MN; oncogene;	
KW	KW	carbonic anhydrase 9; CA9; MN/CA9; carbonic anhydrase IX; CA IX; MN/G250;	
KW	KW	neoplastic tumour; cystostatic; cancer; tumour therapy;	
KW	KW	anti-tumour vaccine.	
OS	OS	Homo sapiens.	
XX	PN	WO2003100029-A2.	
XX	PD	04-DEC-2003.	
XX	XX	22-FEB-2003; 2003WO-US005136.	
Pf	Pf	23-MAY-2002; 2002US-0383068P.	
PR	PR	05-DEC-2002; 2002US-0431499P.	
XX	XX	(FARB ) BAYER CORP.	
PA	PA	(VIRO-) INST VIROLOGY.	
XX	PI	Pastorek J, Pastorekova S, Zatovicova M, Zavada J, Ortova Gut M;	
XX	DR	WPI; 2004-0355136/03.	
DR	DR	P-PSDB; ADG31414.	
XX	PT	New monoclonal antibody generated from MN/CA IX-deficient mice, where the	
PT	PT	antibody binds specifically to human tumor-associated cell adhesion	
PT	PT	protein MN/CA IX or polypeptide, useful for diagnosing, prognosing or	
PT	PT	treating cancer.	
XX	PS	Disclosure; SEQ ID NO 1; 156pp; English.	
XX	CC	This invention relates to a novel monoclonal antibody identified as the	
CC	CC	MN/CA IX specific antibody prepared in knockout mice (CA IX deficient	
CC	CC	mice). Specifically, this antibody is directed towards the MN gene, a	
CC	CC	cellular oncogene known alternatively as carbonic anhydrase 9, CA9 or	
CC	CC	MN/CA9, which encodes the MN protein that is also known as the MN/CA IX	
CC	CC	isoenzyme, carbonic anhydrase IX, CA IX or the MN/G250 protein. The	
CC	CC	present invention describes the generation of this monoclonal antibody,	
CC	CC	and immunoreactive fragments thereof, which are directed against non-	
CC	CC	immunodominant epitopes on the CA IX extracellular domain. As such, this	
CC	CC	antibody can be useful diagnostically as a marker for preneoplastic/	
CC	CC	neoplastic tumors, immunodetection methods and immunotargeting	
CC	CC	approaches. Accordingly, compositions exhibit cytostatic activity and are	
CC	CC	useful in the diagnosis, prognosis and treatment of various cancers	
CC	CC	including breast, bladder or lung cancer, in tumour therapy and in anti-	
CC	CC	tumour vaccination. This polynucleotide sequence is the human MN cDNA of	
CC	CC	the invention.	
XX	SQ	Sequence 1522 BP; 297 A; 458 C; 453 G; 314 T; 0 U; 0 Other;	
 Alignment Scores:			
Pred. No.:	1.85e-130	Length:	1522

Db 973 TTTCAGCGCTACTTCCAAATATGAGGGTCTCTGACTACACCGCCCTGTGCCCAGAGGTGTC 1032  
Qy 341 lletftrpValPheAsnGlnThrValMetLeuSerAlaLysGlnLeuHisThrLeuSer 360  
Db 1033 APTCGACTGTGTTTAAACACAGACAGTGTGCTGAGTGTCTTAACAGCTCCACACCTCTCT 1092  
Qy 361 AspThrLeuTrpGlyProGlyAspSerArgLeuGlnLeuAsnPheArgAlaThrGlnPro 380  
Db 1093 GACACCTGTGGGACCTGTGTACTCTCGCTACAGCTGAACTTCGAGCGACGAGCT 1152  
Qy 381 LeuAsnGlyArgValIleGluAlaSerPheProAlaGlyValAspSerSerProArgAla 400  
Db 1153 TTGAATGGCGAGTGATTGAGGCTCTCTCCCTGCTGGAGTGACAGCAGTCTCTGGGCT 1212  
Qy 401 AlaGluProValGlnLeuAsnSerCysLeuAlaAlaGlyAspIleLeuAlaLeuValPhe 420  
Db 1213 GCTGAGCCAGTCCAGCTGAATTCCTGCTGGCTGTGCTGTGCTGACATCTAGCCCTGTTTTT 1272  
Qy 421 GlyLeuLeuPheAlaValThrSerValAlaPheLeuValGlnMetArgArgGlnHisArg 440  
Db 1273 GGCTCTCTTTTGTGTACACGCGTGGCTCTCTTGTGCAGATGAGAAGGCGACACAGA 1332  
Qy 441 ArgGlyThrLysGlyGlyValSerTyrArgProAlaGluValAlaGluThrGlyAla 459  
Db 1333 AGGGGACCCAAAGGGGTGTGAGTACCAGCCAGCAGAGTAGCCGAGACTGGAGCC 1389

RESULT 9  
ADK41803  
ID ADK41803 standard; DNA; 1522 BP.  
XX  
AC ADK41803;  
XX  
DT 06-MAY-2004 (first entry)  
XX  
DE Human MN gene sequence SeqID1.  
XX  
KW carbonic anhydrase IX; CA IX; precancerous cell; MN; cancerous cell;  
KW human; vertebrate; cytostatic; vaccine; gene therapy;  
KW renal cell carcinoma; breast cancer; colorectal cancer; db; gene.  
XX  
OS Homo sapiens.  
XX  
FH Key Location/Qualifiers  
FT CDS 13..1392  
FT /tag= a  
FT /product= "Human MN protein"  
XX  
XX WO2004/005348-A1.  
XX  
XX 15-JAN-2004.  
XX  
XX 22-FEB-2003; 2003WO-US005137.  
XX  
XX 23-MAY-2002; 2002US-0383068P.  
XX  
XX 05-DEC-2002; 2002US-0431499P.  
XX  
XX (FAR ) BAYER CORP.  
XX  
XX (VIRO-) INST VIROLOGY.  
XX  
XX Zavada J, Pastorekova S, Pastorek J, Zavadova Z;  
XX  
XX WPI; 2004-083500/08.  
XX  
XX P-PSDB; ADK41804.  
XX  
XX New soluble form of the carbonic anhydrase IX (CA IX) protein for  
XX screening, diagnosing or prognosing diseases associated with abnormal  
XX expression of CA IX protein, e.g. renal cell carcinoma, breast cancer or  
XX colorectal cancer.  
XX  
XX Disclosure; SEQ ID NO 1; 159pp; English.  
XX  
XX This invention relates to a novel soluble form of the carbonic anhydrase

CC IX (CA IX) (or MN) protein or CA IX polypeptide which is released from  
CC precancerous and/or cancerous cells of a vertebrate into a body fluid.  
CC The invention may be useful for the development of compounds with a  
CC cytostatic activity or a vaccine whilst the disclosed sequences may be  
CC used for gene therapy. The protein and method are useful for screening,  
CC diagnosing or prognosing diseases associated with abnormal expression of  
CC carbonic anhydrase IX protein, such as precancerous and cancerous  
CC diseases like renal cell carcinoma, breast cancer or colorectal cancer.  
CC The monoclonal antibody may also be used for treating or preventing  
CC precancerous and cancerous diseases. The present sequence is that of the  
CC gene which encodes the human MN protein of the invention.  
XX

Sequence 1522 BP; 297 A; 458 C; 453 G; 314 T; 0 U; 0 Other;

Alignment Scores:  
Pred. No.: 1,85e-130 Length: 1522  
Score: 2424.00 Matches: 459  
Percent Similarity: 100.0% Conservative: 0  
Best Local Similarity: 100.0% Mismatches: 0  
Query Match: 100.0% Indels: 0  
DB: 12 Gaps: 0  
US-09-967-237B-2 (1-459) x ADK41803 (1-1522)  
Qy 1 MetAlaProLeuCysProSerProTrpLeuProLeuLeuLeuProAlaProAlaProGly 20  
Db 13 ATGGCTCCCTGTGCCCCAGCCCTGGCTCCCTCTGTTGATCCCGGCCCTCTCCAGGC 72  
Qy 21 LeuThrValGlnLeuLeuLeuSerLeuLeuLeuLeuMetProValHisProGlnArgLeu 40  
Db 73 CTCACGTGTCAACTGCTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 132  
Qy 41 ProArgMetGlnGluAspSerProLeuGlyGlyGlySerSerGlyGluAspAspProLeu 60  
Db 133 CCCCCGATGCGAGAGGATTCCTCCCTTGGAGAGAGGCTCTTCTGGGAGAGATGACCCACTG 192  
Qy 61 GlyGluGluAspLeuProSerGluGluAspSerProArgGluGluAspProProGlyGlu 80  
Db 193 GCGAGAGGAGGATCTGCCAGTGAAGAGGATTCACCCAGAGAGAGGATCCACCCGAGAG 252  
Qy 81 GluAspLeuProGlyGluGluAspLeuProGlyGluGluAspLeuProGluValLysPro 100  
Db 253 GAGGATCTACCTGCGAGAGAGGATCTACCTGAGAGAGGAGGATCTACCTGAAGTTAAGCCT 312  
Qy 101 LysSerGluGluGlySerLeuLysLeuGluAspLeuProThrValGluAlaProGly 120  
Db 313 AAATCAGAAGAAGAGGGCTCCCTGAAGTTAGAGGATCTACCTACTGTTGAGGCTCTCTGA 372  
Qy 121 AspProGlnGluProGlnAsnAsnAlaHisArgAspLysGluGlyAspAspGlnSerHis 140  
Db 373 GATCTCTAAGAACCCAGATTAATGCCCCACAGGACAAAGAGGGGATGACCCAGAGTCAT 432  
Qy 141 TrpArgTyrGlyGlyAspProProTrpProArgValSerProAlaCysAlaGlyArgPhe 160  
Db 433 TGGCGCTATGAGAGGCGACCCCGCTGGCCCGGGGTGTCCCCAGCCTGCGCGGCGCGCTTC 492  
Qy 161 GlnSerProValAspIleArgProGlnLeuAlaAlaPheCysProAlaLeuArgProLeu 180  
Db 493 CAGTCCCCGGTGGATATCGCCCCCAGCTCGCCGCTCTTTCGCCGCCCTTGGCCCCCTCG 552  
Qy 181 GluLeuLeuGlyPheGlnLeuProProLeuProGluLeuArgLeuArgAsnGlnHis 200  
Db 553 GAACTCTCTGGGCTTCCAGCTCCCGCGCTCCCGAGAACTGCGCTGCGCAACATGSCCAG 612  
Qy 201 SerValGlnLeuThrLeuProProGlyLeuGluMetAlaLeuGlyProGlyArgGluTyr 220  
Db 613 AGTGTGCAACTGACCTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 672  
Qy 221 ArgAlaLeuGlnLeuHisThrGlyAlaAlaGlyArgProGlySerGluHisThr 240  
Db 673 CCGGCTCTGCAGCTGCATCTGCACGTGGGGGCTGACAGGTGCTCCGGGCTCGGAGCAACT 732  
Qy 241 ValGluGlyHisArgPheProAlaGluLeuHisValValHisLeuSerThrAlaPheAla 260

Db 733 GTGGAGGGCCACCGTTTCCCTGCGAGATCCACGTGTTTACCTCAGCACCGCCTTTGGC 792

QY 261 AtqValAspGluAlaLeuGlyArgProGlyGlyLeuAlaValLeuAlaAlaPheLeuGlu 280

Db 793 AGAGTTTGCAGAGCCCTTGGGCGCCCGGAGGCGCTGGCCGTGTGGCCGCTTCTGGAG 852

QY 281 GluGlyProGluGluAenSerAlaTyGluGlnLeuLeuSerArgLeuGluGluLeAla 300

Db 853 GAGGCGCCGAGAGAAACAGTGCCTATGAGCAGTTCGTCTCGTTGGAAGAAATCGCT 912

QY 301 GluGluGlySerGluThrGlnValProGlyLeuAspIleSerAlaLeuProSerAsp 320

Db 913 GAGGAAGGCTCAGAGACTCAGGTCCAGGAGTCGACATATCTGCATCTCGCCCTCTGAC 972

QY 321 PheSerArgTyRPhGlnTyGluGlySerLeuThrProProCysAlaGlnGlyVal 340

Db 973 TTCAGCGCTACTTCCANTATGAGGGGTCTCTGACTACACCGCCCTGTGCCAGGGGTGC 1032

QY 341 IleTrpThrValPheAenGlnThrValMetLeuSerAlaLysGlnLeuHisThrLeuSer 360

Db 1033 ATCTGGACTGTGTTAAACACAGACAGTGTGCTGAGTGTCTAAGCAGCTCCACACCCCTCTCT 1092

QY 361 AspThrLeuTrpGlyProGlyAspSerArgLeuGlnLeuAsnPheArgAlaThrGlnPro 380

Db 1093 GACACCCCTGTGGGAGCCTGTGACTCTCGGCTACAGCTGAACCTCCGAGCGACGACCTT 1152

QY 381 LeuAenGlyArgValIleGluAlaSerPheProAlaGlyValAspSerProArgAla 400

Db 1153 TTGAATGGGCGAGTGATTGAGGCTCTCTTCCCTGCTGGAGTGACAGCTCTCGGGT 1212

QY 401 AlaGluProValGlnLeuAenSerCysLeuAlaAlaGlyAspIleLeuAlaLeuValPhe 420

Db 1213 GCTGAGCCAGTCCAGCTGAATTCCTCGCTGGCTGCTGGTGACATCTAGCCCTGTGTTT 1272

QY 421 GlyLeuLeuPheAlaValThrSerValAlaPheLeuValGlnMetArgArgGlnHisArg 440

Db 1273 GGCTCTCTTTTGTCTGTCACAGCGTGGGTCTCTTGTGAGATGAGAAGGACGACAGA 1332

QY 441 ArgGlyThrLysGlyGlyValSerTyArgProAlaGluValAlaGluThrGlyAla 459

Db 1333 AGGGGAAACCAAGGGGTGTGACTACCGCCGACGAGAGGTAGCCGAGCTGGAGCC 1389

RESULT 10

ADZ64594

ID ADZ64594 standard; cDNA; 1522 BP.

XX AC ADZ64594;

XX DT 14-JUL-2005 (first entry)

XX DE Human MN/CA IX cDNA.

XX KW Prognosis; oncogene; neoplasm; MN/CA IX; stomach tumor;

XX KW gallbladder disease; biliary cancer; biliary tumor; testis tumor;

XX KW ovary tumor; basal cell carcinoma; central nervous system tumor; gene;

XX OS Homo sapiens.

XX FH Key Location/Qualifiers

XX FT CDS 13..1392

FT /\*tag= b

FT /product= "Human MN/CA IX"

FT sig\_peptide 13..123

FT /\*tag= a

FT mat\_peptide 124..1389

FT /\*tag= c

FT /product= "Mature MN/CA IX"

XX PN WO2005037083-A2.

XX XX 28-APR-2005.

XX 18-OCT-2004; 2004WO-US034573.

XX 16-OCT-2003; 2003US-0511832P.

XX (FARB ) BAYER HEALTHCARE.

PA (VIRO-) INST VIROLOGY.

XX Ebert M, Rocken C, Pastorekova S, Zavada J, Pastorek J;

PI WPI: 2005-315588/32.

XX P-PSDB; ADZ64595.

DR Prognosticating preneoplastic/neoplastic diseases of the breast,

PT gallbladder, biliary ducts, testis or ovaries, by detecting MN/CA9 gene

PT expression products, useful in medical genetics, biochemical engineering

PT and immunochemistry.

XX Claim 1; SEQ ID NO 1; 83pp; English.

XX The invention relates to a method of prognosticating a

CC preneoplastic/neoplastic disease afflicting a subject vertebrate, where

CC the disease affects a tissue that normally expresses an MN/CA IX protein

CC but loses or has significantly reduced MN/CA IX expression upon

CC carcinogenesis, comprising detecting an MN/CA9 gene expression product in

CC a sample having preneoplastic/neoplastic tissue taken from the

CC vertebrate, quantitating the level of the MN/CA9 gene expression product

CC in the sample, comparing the level of MN/CA9 gene expression product to

CC the average level of MN/CA9 gene expression product in comparable samples

CC taken from vertebrates afflicted by the same preneoplastic/neoplastic

CC disease as the subject vertebrate, and determining that the subject

CC vertebrate has a poorer prognosis if the level of MN/CA9 gene expression

CC product quantitated is higher than the average level of MN/CA9 gene

CC expression product in the comparable samples. The methods and

CC compositions of the present invention are useful in medical genetics,

CC biochemical engineering, immunochemistry and oncology, in particular for

CC prognosticating preneoplastic/neoplastic diseases of gastric mucosa,

CC gallbladder, biliary ducts, ductal cells of duodenal glands, testis

CC including ductular efferens and rete testis, ovary including surface

CC coelomic epithelium and rete ovary, basal cells of hair follicles and

CC central nervous system choroid plexus. This sequence represents cDNA

CC encoding the human MN/CA IX polypeptide of the invention.

XX SQ Sequence 1522 BP; 297 A; 458 C; 453 G; 314 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 1-85e-130 Length: 1522

Score: 2424.00 Matches: 459

Percent Similarity: 100.0% Conservatives: 0

Best Local Similarity: 100.0% Mismatches: 0

Query Match: 100.0% Indels: 0

DB: 14 Gaps: 0

US-09-967-237B-2 (1-459) x ADZ64594 (1-1522)

QY 1 MetAlaProLeuGlyProSerProTrpLeuProLeuLeuLeuProAlaProGly 20

Db 13 ATGGCTCCCTGTGCCAGCGCCCTGCTCCCTCTGTGATCCCGCCCTGTCTCCAGGC 72

QY 21 LeuThrValGlnLeuLeuSerLeuLeuLeuLeuMetProValHisProGlnArgLeu 40

Db 73 CTCATGTGCAACTGCTGTGTCACGTGCTGTCTGATGCCGTCTATCCCGAGGTTG 132

QY 41 ProArgMetGlnGluAspSerProLeuGlyGlySerSerGlyGluAspAspProLeu 60

Db 133 CCCCCGATGCAGGAGATTCCCCCTTGGGAGGAGGCTCTTCTGGGGAAGATGACCCACTG 192

QY 61 GlyGluGluAspLeuProSerGluGluAspSerProArgGluGluAspProProGlyGlu 80

Db 193 GGCAGAGGAGTCTGCCAGTAGAGGAGTTCACCCAGAGAGAGGATCCACCCGAGAG 252

QY 81 GluAspLeuProGlyGluGluAspLeuProGlyGluGluAspLeuProGlyValLysPro 100

Db 253 GAGGATCTACCTGGAGGAGGATCTACCTGGAGGAGGATCTACCTGAAGTTAAGCCT 312  
Qy 101 LysSerGluGluGluGlySerLeuLysLeuGluAspLeuProThrValGluAlaProGly 120  
Db 313 AAATCAGAGAGAGGCTCCCTGAAGTTAGAGATCTACCTACTGTTAGGCTCTCTGGA 372  
Qy 121 AspProGlnGluProGlnAenAenAlaHisArgAspLysGluGlyAspAspGlnSerHis 140  
Db 373 GATCTCTAGAACCCAGATATATCCACAGGACAAAGAGGGGATGACACAGATCAT 432  
Qy 141 TrpArgTyrGlyGlyAspProProTrpProArgValSerProAlaCysAlaGlyArgPhe 160  
Db 433 TGGCGCTATGAGGCGACCCGCTGCGCCCGGGTGTCCACAGCTGCGGGCGCGCTTC 492  
Qy 161 GlnSerProValAspIleArgProGlnLeuAlaPheCysProAlaLeuArgProLeu 180  
Db 493 CAGTCCCGGTGATATCCGCCCCAGCTCGCCGCTTCTGCCCCGGCCCTGCGCCCCCTG 552  
Qy 181 GluLeuLeuGlyPheGlnLeuProProLeuProGluLeuArgLeuArgAsnAnGlyHis 200  
Db 553 GAATCTCTGGCTTCCAGCTCCCGCGCTCCAGNACTGCGCTCGCGCAATATGCCAC 612  
Qy 201 SerValGlnLeuThrLeuProProGlyLeuGluMetAlaLeuGlyProGlyArgGluTyr 220  
Db 613 AGTGTCAACTGACCTGCTCTCTGGCTAGAGATGGCTCTGGGCTCCCGGGCGGAGTAC 672  
Qy 221 ArgAlaLeuGlnLeuHisLeuHisTyrGlyValAlaGlyValArgProGlySerGluHisThr 240  
Db 673 CGGGCTCTGAGCTGCAATCTGGGGGCTGCGAGGTGCTCGGGCTCGGAGCACACT 732  
Qy 241 ValGluGlyHisArgPheProAlaGluIleHisValHisLeuSerThrAlaPheAla 260  
Db 733 GTGGAGGCCACCGTTTCCCTCGAGATCCAGTGTTCCTCAGCACCGCTTTGCC 792  
Qy 261 ArgValAspGluAlaLeuGlyArgProGlyGlyLeuAlaValLeuAlaAlaPheLeuGlu 280  
Db 793 AGAGTTGACGAGGCTTTGGGGCGCCGGGAGGCTGGCGCTGTGGCGGCTTTCTGGAG 852  
Qy 281 GluGlyProGluGluAanSerAlaTyrGluGlnLeuLeuSerArgLeuGluIleAla 300  
Db 853 GAGGGCCCGGAGAAACAGTGCTATGAGCAGTTGCTGCTCGCTTGGAGAAATCGCT 912  
Qy 301 GluGluGlySerGluThrGlnValProGlyLeuAspIleSerAlaLeuLeuProSerAsp 320  
Db 913 GAGGAGGCTCAGAGACTCAGTCCAGGACTGGACATATCTGCATCTCTGCCCTCTGAC 972  
Qy 321 PheSerArgTyrPheGlnTyrGluGlySerLeuThrThrProProCysAlaGlnGlyVal 340  
Db 973 TTCAGCGCTACTTCCAAATATGAGGGTCTCTGACTACACCGGCTGTGCGCGGCTGTC 1032  
Qy 341 IleThrThrValPheAanGlnThrValMetLeuSerAlaLysGlnLeuHisThrLeuSer 360  
Db 1033 ATCTGGACTGTGTTTAAACACAGACAGTATGCTGAGTCTAAGCAGCTCCACACCTCTCT 1092  
Qy 361 AspThrLeuTrpGlyProGlyAspSerArgLeuGlnLeuAanPheArgAlaThrGlnPro 380  
Db 1093 GACACCTGTGGGACCTGTGACTCTCGGCTACAGTGAATCTCCGAGCAGCAGCCT 1152  
Qy 381 LeuAanGlyArgValIleGluAlaSerPheProAlaGlyValAspSerSerProArgAla 400  
Db 1153 TTGAATGGCGAGTGATTTGAGGCTCTTCCCTGCTGAGTGGAGTGGACAGCAGTCTCGGGCT 1212  
Qy 401 AlaGluProValGlnLeuAanSerCysLeuAlaAlaGlyAspIleLeuAlaLeuValPhe 420  
Db 1213 GCTGAGCAGTCCACAGCTGAAATCTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1272  
Qy 421 GlyLeuLeuPheAlaValThrSerValAlaPheLeuValGlnMetArgArgGlnHisArg 440  
Db 1273 GGCTCTCTTTTGTCTGTCACAGGCTCGGCTCTCTGTCAGATGAGAGCGACAGAGA 1332  
Qy 441 ArgGlyThrLysGlyGlyValSerTyrArgProAlaGluValAlaGluThrGlyAla 459  
Db 1333 AGGGGAACCAAGGGGTGTGAGCTACCGCCAGCAGAGGTAGCCGAGACTGGAGCC 1389

## RESULT 11

ABL64779  
ID ABL64779 standard; DNA; 1552 BP.  
XX  
AC ABL64779;  
XX  
DT 15-MAY-2002 (first entry)  
XX  
DE Lung cancer related gene sequence SEQ ID NO:3116.  
XX  
KW Human; cancer; colon; breast; kidney; thyroid;  
KW stomach; lung; prostate; pancreas; carcinoma; antitumour; cancerous;  
KW cytostatic; gene therapy; antineoplastic; Wilms' tumour; adenocarcinoma;  
KW gene; ds.  
XX  
OS Homo sapiens.  
XX  
PN WO200194629-A2.  
XX  
PD 13-DEC-2001.  
XX  
PF 30-MAY-2001; 2001WO-US010838.  
XX  
PR 05-JUN-2000; 2000US-0209473P.  
PR 05-JUN-2000; 2000US-0209531P.  
PR 18-SEP-2000; 2000US-0233133P.  
PR 18-SEP-2000; 2000US-0233617P.  
PR 20-SEP-2000; 2000US-0234009P.  
PR 20-SEP-2000; 2000US-0234034P.  
PR 20-SEP-2000; 2000US-0234052P.  
PR 22-SEP-2000; 2000US-0234509P.  
PR 22-SEP-2000; 2000US-0234567P.  
PR 25-SEP-2000; 2000US-0234923P.  
PR 25-SEP-2000; 2000US-0234924P.  
PR 25-SEP-2000; 2000US-0235077P.  
PR 25-SEP-2000; 2000US-0235082P.  
PR 25-SEP-2000; 2000US-0235134P.  
PR 25-SEP-2000; 2000US-0235280P.  
PR 26-SEP-2000; 2000US-0235637P.  
PR 26-SEP-2000; 2000US-0235638P.  
PR 27-SEP-2000; 2000US-0235711P.  
PR 27-SEP-2000; 2000US-0235720P.  
PR 27-SEP-2000; 2000US-0235840P.  
PR 27-SEP-2000; 2000US-0235863P.  
PR 28-SEP-2000; 2000US-0236028P.  
PR 28-SEP-2000; 2000US-0236032P.  
PR 28-SEP-2000; 2000US-0236033P.  
PR 28-SEP-2000; 2000US-0236034P.  
PR 28-SEP-2000; 2000US-0236109P.  
PR 28-SEP-2000; 2000US-0236111P.  
PR 29-SEP-2000; 2000US-0236842P.  
PR 29-SEP-2000; 2000US-0236891P.  
PR 02-OCT-2000; 2000US-0237172P.  
PR 02-OCT-2000; 2000US-0237173P.  
PR 02-OCT-2000; 2000US-0237278P.  
PR 02-OCT-2000; 2000US-0237294P.  
PR 02-OCT-2000; 2000US-0237295P.  
PR 02-OCT-2000; 2000US-0237316P.  
PR 03-OCT-2000; 2000US-0237425P.  
PR 03-OCT-2000; 2000US-0237598P.  
PR 03-OCT-2000; 2000US-0237604P.  
PR 03-OCT-2000; 2000US-0237606P.  
PR 03-OCT-2000; 2000US-0237608P.  
PR 01-NOV-2000; 2000US-0244867P.  
PR 01-NOV-2000; 2000US-0244868P.  
XX  
XX (AVAL-) AVALON PHARM.  
XX  
XX Young PE, Augustus M, Carter KC, Ebner R, Endress G, Horrigan S;  
XX Soppet DR, Weaver Z;  
XX  
XX WPI; 2002-188264/24.  
DR



PN	W0200194629-A2.	CC	properties of the agent. M1 can be used in the treatment of cancer such
XX		CC	as colon, breast, stomach, lung, thyroid, oesophageal, ovarian, kidney,
PD	13-DEC-2001.	CC	prostate or pancreatic cancer, adenocarcinoma, carcinoma, clear cell
XX		CC	cancer, infiltrating ductal cancer, infiltrating lobular cancer, squamous
PP	30-MAY-2001; 2001WO-US010838.	CC	cell carcinoma, neuroendocrine carcinoma, papillary carcinoma and Wilm's
XX		XX	tumour
PR	05-JUN-2000; 2000US-0209473P.	SQ	Sequence 1552 BP; 302 A; 471 C; 461 G; 318 T; 0 U; 0 Other;
PR	05-JUN-2000; 2000US-0209531P.	Alignment Scores:	
PR	18-SEP-2000; 2000US-0233133P.	Pred. No.:	1.88e-130 Length: 1552
PR	18-SEP-2000; 2000US-0233133P.	Score:	2424.00 Matches: 459
PR	20-SEP-2000; 2000US-0234009P.	Percent Similarity:	100.0% Conservative: 0
PR	20-SEP-2000; 2000US-0234009P.	Best Local Similarity:	100.0% Mismatches: 0
PR	20-SEP-2000; 2000US-0234034P.	Query Match:	100.0% Indels: 0
PR	22-SEP-2000; 2000US-0234509P.	DB:	6 Gaps: 0
PR	22-SEP-2000; 2000US-0234567P.		
PR	25-SEP-2000; 2000US-0234923P.	US-09-967-237B-2 (1-459) x ABL62179 (1-1552)	
PR	25-SEP-2000; 2000US-0234924P.		
PR	25-SEP-2000; 2000US-0235077P.		
PR	25-SEP-2000; 2000US-0235082P.		
PR	25-SEP-2000; 2000US-0235134P.		
PR	25-SEP-2000; 2000US-0235280P.		
PR	25-SEP-2000; 2000US-0235280P.		
PR	26-SEP-2000; 2000US-0235637P.		
PR	26-SEP-2000; 2000US-0235638P.		
PR	27-SEP-2000; 2000US-0235711P.		
PR	27-SEP-2000; 2000US-0235720P.		
PR	27-SEP-2000; 2000US-0235840P.		
PR	27-SEP-2000; 2000US-0235863P.		
PR	28-SEP-2000; 2000US-0236028P.		
PR	28-SEP-2000; 2000US-0236032P.		
PR	28-SEP-2000; 2000US-0236033P.		
PR	28-SEP-2000; 2000US-0236034P.		
PR	28-SEP-2000; 2000US-0236109P.		
PR	28-SEP-2000; 2000US-0236111P.		
PR	29-SEP-2000; 2000US-0236842P.		
PR	29-SEP-2000; 2000US-0236891P.		
PR	02-OCT-2000; 2000US-0237172P.		
PR	02-OCT-2000; 2000US-0237173P.		
PR	02-OCT-2000; 2000US-0237278P.		
PR	02-OCT-2000; 2000US-0237294P.		
PR	02-OCT-2000; 2000US-0237295P.		
PR	03-OCT-2000; 2000US-0237316P.		
PR	03-OCT-2000; 2000US-0237425P.		
PR	03-OCT-2000; 2000US-0237598P.		
PR	03-OCT-2000; 2000US-0237604P.		
PR	03-OCT-2000; 2000US-0237606P.		
PR	03-OCT-2000; 2000US-0237608P.		
PR	01-NOV-2000; 2000US-0244867P.		
PR	01-NOV-2000; 2000US-0245084P.		
XX	(AVAL-) AVALON PHARM.		
XX	Young PE, Augustus M, Carter KC, Ebner R, Endress G, Horrigan S;		
PI	Soppet DR, Weaver Z;		
PI	WPI; 2002-188264/24.		
XX			
DR			
XX			
PT	Screening for anti-neoplastic agent involves exposing cells to a chemical		
PT	agent to be tested for anti-neoplastic activity, and determining a change		
PT	in expression of a gene of a signature gene set.		
XX			
XX	Claim 1; SEQ ID NO 516; 44pp; English.		
XX			
CC	The present invention describes a method (M1) for screening for an anti-		
CC	neoplastic agent. The method involves exposing cells to a chemical agent		
CC	to be tested for anti-neoplastic activity, determining a change in		
CC	expression of at least one gene (I) of a signature gene set, where (I)		
CC	comprises a sequence (S) selected from 8447 sequences (given in ABL6164		
CC	to ABL70110), or is at least 95% identical to (S), where a change in		
CC	expression is indicative of anti-neoplastic activity. (I) has cytostatic		
CC	activity and can be used in gene therapy. M1 can be used for screening an		
CC	anti-neoplastic agent, and can be used for producing a product which is		
CC	the data collected with respect to the anti-neoplastic agent as a result		
CC	of M1, and the data is sufficient to convey the chemical structure and/or		

```
Db      823  AGAGTTGACGAGGCTTGGGGGCGCCGGGAGGCGCTGGCCGTGTGGCCGCTTTCTGGAG 882
Qy      281  GluGlyProGluGluAenSerAlaTyrGluGlnLeuLeuSerArgLeuGluGluIleAla 300
      |||||
Db      883  GAGGCGCCGGAAGAAACAGTGCTATGAGCAGTTGCTGCTCGCTTGGGAAGAAATCGCT 942
Qy      301  GluGluGlySerGlnThrGlnValProGlyLeuAspIleSerAlaLeuProSerAsp 320
      |||||
Db      943  GAGGAAGGCTCAGAGACTCAGGTCCTCCAGGACTGGACATATCTGCATCTCCCTGCCCTCTGAC 1002
Qy      321  PheSerArgTyrPheGlnTyrGluGlySerLeuThrProProCysAlaGlnGlyVal 340
      |||||
Db      1003  TTGAGCGCTACTTCCATATGAGGGTCTTCTGACTACACCGCCCTGTGCCCAAGGTGTC 1062
Qy      341  IleTrpThrValPheAenGlnThrValMetLeuSerAlaLysGlnLeuHisThrLeuSer 360
      |||||
Db      1063  ATCTGGACTGTGTTTAACACAGACAGTGTCTGAGTGTAAAGCAGCTCCACACCCCTCTCT 1122
Qy      361  AspThrLeuTrpGlyProGlyAspSerArgLeuGlnLeuAenPheArgAlaThrGlnPro 380
      |||||
Db      1123  GACACCCCTGTGGGACCTGGTGACTCTCGGCTACAGCTGAACCTTCCGAGCGACGACCT 1182
Qy      381  LeuAsnGlyArgValIleGluAlaSerPheProAlaGlyValAspSerProArgAla 400
      |||||
Db      1183  TTGAATGGGCGAGTGATTGAGGCTCTTCCCTGCTGGAGTGGACAGCAGTCTCTCGGCT 1242
Qy      401  AlaGluProValGlnLeuAenSerCysLeuAlaAlaGlyAspIleLeuAlaLeuValPhe 420
      |||||
Db      1243  GCTGAGCCAGTCCAGCTGAATCTCGCTGGCTGCTGTGACATCTAGCCCTGTGTTTTT 1302
Qy      421  GlyLeuLeuPheAlaValThrSerValAlaPheLeuValGlnMetArgGlnHisArg 440
      |||||
Db      1303  GGCTCTCTTTTCTGTGCACAGCGTGGCGTTCCTTTGTGCAGATGAGAAGCGACAGACA 1362
Qy      441  ArgGlyThrLysGlyGlyValSerTyrArgProAlaGluValAlaGluThrGlyAla 459
      |||||
Db      1363  AGGGGAACCAAGGGGTGTGAGCTACCGCCGACGAGGTAGCCGAGACTGGAGCC 1419

RESULT 13
ABL65416
ID      ABL65416 standard; DNA; 1552 BP.
XX
AC      ABL65416;
XX
DT      15-MAY-2002 (first entry)
XX
DE      Lung cancer related gene sequence SEQ ID NO:3753.
XX
KW      Human; cancer; colon; breast; ovary; oesophagus; kidney; thyroid;
KW      stomach; lung; prostate; pancreas; carcinoma; antitumour; cancerous;
KW      cytostatic; gene therapy; antineoplastic; Wilm's tumour; adenocarcinoma;
KW      gene; ds.
XX
OS      Homo sapiens.
XX
PN      WO200194629-A2.
XX
PD      13-DEC-2001.
XX
PF      30-MAY-2001; 2001WO-US010838.
XX
PR      05-JUN-2000; 2000US-0209473P.
PR      05-JUN-2000; 2000US-0209531P.
PR      18-SEP-2000; 2000US-0231333P.
PR      18-SEP-2000; 2000US-0233617P.
PR      20-SEP-2000; 2000US-0234009P.
PR      20-SEP-2000; 2000US-0234034P.
PR      20-SEP-2000; 2000US-0234052P.
PR      22-SEP-2000; 2000US-0234509P.
PR      22-SEP-2000; 2000US-0234567P.
PR      25-SEP-2000; 2000US-0234923P.
PR      25-SEP-2000; 2000US-0234924P.
PR      25-SEP-2000; 2000US-0235077P.
```

```
PR      25-SEP-2000; 2000US-0235082P.
PR      25-SEP-2000; 2000US-0235134P.
PR      25-SEP-2000; 2000US-0235280P.
PR      26-SEP-2000; 2000US-0235637P.
PR      26-SEP-2000; 2000US-0235638P.
PR      27-SEP-2000; 2000US-0235711P.
PR      27-SEP-2000; 2000US-0235720P.
PR      27-SEP-2000; 2000US-0235840P.
PR      27-SEP-2000; 2000US-0235863P.
PR      28-SEP-2000; 2000US-0236028P.
PR      28-SEP-2000; 2000US-0236032P.
PR      28-SEP-2000; 2000US-0236033P.
PR      28-SEP-2000; 2000US-0236034P.
PR      28-SEP-2000; 2000US-0236109P.
PR      28-SEP-2000; 2000US-0236111P.
PR      29-SEP-2000; 2000US-0236842P.
PR      29-SEP-2000; 2000US-0236891P.
PR      02-OCT-2000; 2000US-0237172P.
PR      02-OCT-2000; 2000US-0237173P.
PR      02-OCT-2000; 2000US-0237278P.
PR      02-OCT-2000; 2000US-0237294P.
PR      02-OCT-2000; 2000US-0237295P.
PR      02-OCT-2000; 2000US-0237316P.
PR      03-OCT-2000; 2000US-0237425P.
PR      03-OCT-2000; 2000US-0237598P.
PR      03-OCT-2000; 2000US-0237604P.
PR      03-OCT-2000; 2000US-0237606P.
PR      03-OCT-2000; 2000US-0237608P.
PR      01-NOV-2000; 2000US-0244867P.
PR      01-NOV-2000; 2000US-0245084P.
XX
PA      (AVAL-) AVALON PHARM.
XX
PI      Young PR, Augustus M, Carter KC, Ebner R, Endress G, Horrigan S;
PI      Soppet DR, Weaver Z;
XX
DR      WPI; 2002-188264/24.
XX
PT      Screening for anti-neoplastic agent involves exposing cells to a chemical
PT      agent to be tested for anti-neoplastic activity, and determining a change
PT      in expression of a gene of a signature gene set.
XX
PS      Claim 1; SEQ ID NO 3753; 44pp; English.
XX
CC      The present invention describes a method (M1) for screening for an anti-
CC      neoplastic agent. The method involves exposing cells to a chemical agent
CC      to be tested for anti-neoplastic activity, determining a change in
CC      expression of at least one gene (I) of a signature gene set, where (I)
CC      comprises a sequence (S) selected from 8447 sequences (given in ABL61664
CC      to ABL70110), or is at least 95% identical to (S), where a change in
CC      expression is indicative of anti-neoplastic activity. (I) has cytostatic
CC      activity and can be used in gene therapy. M1 can be used for screening an
CC      anti-neoplastic agent, and can be used for producing a product which is
CC      the data collected with respect to the anti-neoplastic agent as a result
CC      of M1, and the data is sufficient to convey the chemical structure and/or
CC      properties of the agent. M1 can be used in the treatment of cancer such
CC      as colon, breast, stomach, lung, thyroid, oesophageal, ovarian, kidney,
CC      prostate or pancreatic cancer, adenocarcinoma, carcinoma, clear cell
CC      cancer, infiltrating ductal cancer, infiltrating lobular cancer, squamous
CC      cell carcinoma, neuroendocrine carcinoma, papillary carcinoma and Wilm's
CC      tumour
XX
SQ      Sequence 1552 BP; 302 A; 471 C; 461 G; 318 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:      1.88e-130      Length:      1552
Score:          2424.00      Matches:      459
Percent Similarity: 100.0%      Conservatives: 0
Best Local Similarity: 100.0%      Mismatches: 0
Query Match:    100.0%      Indels:      0
DB:            6      Gaps:      0

US-09-967-237B-2 (1-459) x ABL65416 (1-1552)
```







```
QY 341 IleTrrThrValPheAsnGlnThrValMetLeuSerAlaLysGlnLeuHisThrLeuSer 360
Db |||||
1063 ATCTGGACTGTGTTTAACACAGACAGTGATGCTGAGTGCTAAGCAGCTCCACACCCCTCTCT 1122
QY 361 AspThrLeuTrrGlyProGlyAspSerArgLeuGlnLeuAsnPheArgAlaThrGlnPro 380
Db |||||
1123 GACACCCCTGTGGGACCTGTGTGACTCTCGGCTACAGCTGAACCTCCGAGCGACGAGCCCT 1182
QY 381 LeuAsnGlyArgValIleGluAlaSerPheProAlaGlyValAspSerSerProArgAla 400
Db |||||
1183 TTGAATGGGCGAGTGATTGAGGCTCTCTCCCTGCTGGAGTGGACAGCAGTCCTCGGGCT 1242
QY 401 AlaGluProValGlnLeuAsnSerCysLeuAlaAlaGlyAspIleLeuAlaLeuValPhe 420
Db |||||
1243 GCTGAGCCAGTCCAGCTGAATTCTGCTGGCTGGTGTGACATCCTAGCCCTGGTTTTT 1302
QY 421 GlyLeuLeuPheAlaValThrSerValAlaPheLeuValGlnMetArgArgGlnHisArg 440
Db |||||
1303 GGCCTCCTTTTGTGTGTCACCGGTGGCGTTCTTGTGCAGATGAGAAGGCAGCACAGA 1362
QY 441 ArgGlyThrLysGlyGlyValSerTyrArgProAlaGluValAlaGluThrGlyAla 459
Db |||||
1363 AGGGGAACCAAGGGGGTGTGAGCTACCGCCAGCAGAGGTAGCCGAGACTGGAGCC 1419
```

Search completed: February 17, 2006, 14:47:57  
Job time : 864 secs

GenCore version 5.1.7  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: February 17, 2006, 14:48:07 ; Search time 5571 Seconds

(without alignments)  
3854.832 Million cell updates/sec

Title: US-09-967-237B-2

Perfect score: 2424

Sequence: 1 MAPLCSPWLPPLIPAPAG.....RRGTGKGVSRPAEVAETGA 459

Scoring table:

BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 41078325 seqs, 23393541228 residues

Total number of hits satisfying chosen parameters: 82156650

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh  
-Q=/abs/ABSSWEB.spool/US09967237/runat\_15022006.182244\_28001/app\_query.fasta\_1  
-DB=EST -OPMT=fastcap -SUFFIX=rst -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0  
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45  
-DOALIGN=200 -THR\_SCORE=pcr -THR\_MAX=100 -THR\_MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=ptc -NORM=ext -HEADSIZE=500 -MINLEN=0 -MAXLEN=2000000000 -HOST=abs03h  
-USER=US09967237 @CGN 1.1 5315 @runat\_15022006.182244\_28001 -NCPU=6 -ICPU=3  
-NO MMAP -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG -DEV TIMEOUT=120  
-WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7  
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

EST:\*

1: gb\_est1:\*  
2: gb\_est2:\*  
3: gb\_est3:\*  
4: gb\_est4:\*  
5: gb\_est5:\*  
6: gb\_est6:\*  
7: gb\_est7:\*  
8: gb\_est8:\*  
9: gb\_est9:\*  
10: gb\_gss2:\*  
11: gb\_gss3:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2420	99.8	1469	4	CR616345 full-leng
2	2420	99.8	1523	4	CR620502 full-leng
3	2420	99.8	1541	4	CR597234 full-leng
4	2361.5	97.4	1492	4	CR590646 full-leng
5	2177	89.8	1380	11	DQ049266 Homo sapi
6	1926	79.5	1380	11	DQ049267 Pan trogl
7	1790	73.8	1084	1	AL542336

8	1567	64.6	1070	1	AL554705
9	1556.5	64.2	1009	1	AL558378
10	1501.5	61.9	1017	5	AX401186
11	1472	60.7	916	5	AX423970
12	1452.5	59.9	971	7	CO579387
13	1401	57.8	927	1	AL555184
14	1370	56.1	830	8	DR432523
15	1360	56.5	887	1	AJ924825
16	1332	55.0	1071	1	AL554665
17	1319	54.4	833	1	DR432759
18	1304.5	53.8	874	2	EG386425
19	1289.5	53.2	815	8	DR432519
20	1260	52.0	811	8	DR432758
21	1223	50.5	693	8	DR433331
22	1211	50.0	836	5	AX383092
23	1160	47.9	692	8	DR432847
24	1130	46.6	740	7	CK971835
25	1122.5	46.3	687	8	DR432755
26	1121	46.2	675	8	DR432851
27	1121	46.2	1017	1	AL580216
28	1115	46.0	994	1	AL577748
29	1114	46.0	668	2	EG824243
30	1113	45.9	818	2	BE548062
31	1094	45.1	975	1	AJ924826
32	1092	45.0	782	7	CK969500
33	1090	45.0	961	5	AX423969
34	1075	44.3	722	7	CK978804
35	1064.5	43.9	746	2	BI223232
36	1060.5	43.8	806	2	BF344769
37	1041	42.9	589	3	BM790508
38	1041	42.9	813	7	CO734371
39	1007.5	41.6	962	7	CV675676
40	1003	41.4	663	8	DR432756
41	1002.5	41.4	736	2	BE543633
42	998.5	41.2	666	2	DR433328
43	989	40.8	830	7	CK849799
44	986.5	40.7	945	6	CF579370
45	965	39.8	715	7	CK963805

#### ALIGNMENTS

RESULT 1	CR616345	1469 bp	mRNA	linear	HTC 21-JUL-2004
LOCUS	full-length cDNA clone CS0DA003YB12 of Neuroblastoma of Homo sapiens (human).				
DEFINITION	CR616345				
ACCESSION	CR616345.1	GI:50497152			
VERSION	HTC; CNSLIT cdNA.				
KEYWORDS	Homo sapiens				
SOURCE	Homo sapiens				
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.				
REFERENCE	1 (bases 1 to 1469)				
AUTHORS	Li, W.B., Gruber, C., Jesse, J. and Polayes, D.				
TITLE	Full-length cDNA libraries and normalization				
JOURNAL	Unpublished				
REMARK	Contact: Feng Liang Email: fliang@lifetech.com URL: http://fulllength.invitrogen.com/ Invitrogen Corporation 1600 Paraday Avenue				
REFERENCE	2 (bases 1 to 1469)				
AUTHORS	Genoscope.				
TITLE	Direct Submission				
JOURNAL	Submitted (20-JUL-2004) Genoscope - Centre National de Sequencage : BP 191 91006 EVRY cedex - FRANCE (E-mail: secre@genoscope.cns.fr)				
COMMENT	Web : www.genoscope.cns.fr 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime ends enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized. Library was constructed by Life Technologies, a division of Invitrogen.				



/mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clones="CS0DK007YK10"  
 /issue\_type="Hela cells Cot 25-normalized"  
 /plasmid="pCMVSPORT\_6"

## ORIGIN

## Alignment Scores:

Pred. No.: 3,07e-185 Length: 1523  
 Score: 2420.00 Matches: 458  
 Percent Similarity: 100.0% Conservative: 1  
 Best Local Similarity: 99.8% Mismatches: 0  
 Query Match: 99.8% Indels: 0  
 DB: 4 Gaps: 0

US-09-967-237B-2 (1-459) x CR620502 (1-1523)

Qy 1 MetAlaProLeuCySProSerProTrpLeuProLeuLeuLeuProAlaProAlaProGly 20  
 Db 32 ATGGCTCCCTGTGCCCGCCCTGGCTCCCTCTGTGATCCCGCCCTGTCTCCAGGC 91  
 Qy 21 LeuThrValGlnLeuLeuLeuSerLeuLeuLeuMetProValHisProGlnArgLeu 40  
 Db 92 CTCACTGTCAACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 151  
 Qy 41 ProArgMetGlnGluAspSerProLeuGlyGlyGlySerSerGlyGluAspProLeu 60  
 Db 152 CCCCGAGTGCAGGAGGATTCCTCCCTTGGAGAGGCTCTTCTGGGGAGATGACCCACTG 211  
 Qy 61 GlyGluGluAspLeuProSerGluGluAspSerProArgGluGluAspProProGlyGlu 80  
 Db 212 GCGGAGGAGNTCTGCCAGTAGAGAGATTCACCCAGAGAGGAGATCCACCCGAGAG 271  
 Qy 81 GluAspLeuProGlyGluGluAspLeuProGlyGluGluAspLeuProGluValLysPro 100  
 Db 272 GAGGATCTACTGAGAGGAGGATCTACTCTGGAGAGGAGGATCTACTGAGAGGATCT 331  
 Qy 101 LysSerGluGluGlySerLeuLysLeuGluAspLeuProThrValGluAlaProGly 120  
 Db 332 AAATCAAGAAGAGGGCTCCCTCAAGTTAGAGATCTACTGTTGAGGCTCTCTGGA 391  
 Qy 121 AspProGlnGluProGlnAsnAlaHisArgAspLysGluGlyAspAspGlnSerHis 140  
 Db 392 GATCTCTCAAGAACCCAGATAATGCCACAGGACAAAGAGGGATGACACAGATCAT 451  
 Qy 141 TrpArgTyrGlyAspProTrpProArgValSerProAlaCysAlaGlyArgPhe 160  
 Db 452 TGGCGCTATGAGGCGACCCCGCTGCGCGCGGTGTCCTCCAGCGCTGCGCGCGCGCTTC 511  
 Qy 161 GlnSerProValAspLysArgProGlnLeuAlaAlaPheCysProAlaLeuAspProLeu 180  
 Db 512 CAGTCCCGGFGGATATCCCGCCCTGCTGCGCGCTCTGCGCGCGCTGCGCGCGCTG 571  
 Qy 181 GluLeuLeuGlyPheGlnLeuProLeuProGluLeuArgLeuArgAsnAsnGlyHis 200  
 Db 572 GAATCTCTGGGCTTCCAGCTCCCGCGCTCCAGACTGCGGCTGCGGACAAATGGCCAC 631  
 Qy 201 SerValGlnLeuThrLeuProProGlyLeuGluMetAlaLeuGlyProGlyArgGlyTyr 220  
 Db 632 AGTGTCAACTGACCTGCTCTCTGGCTAGAGATGCTCTGGGTCCCGGCGCGGAGTAC 691  
 Qy 221 ArgAlaLeuGlnLeuHisLeuHisTrpGlyAlaAlaGlyArgProGlySerGluHisThr 240  
 Db 692 CGGGCTCTGCAGCTGCATCTGCTGCGGGGCTCGAGTCTGCGGGCTCGGAGCAACT 751  
 Qy 241 ValGluGlyHisArgPheProAlaGluLeuHisValValHisLeuSerThrAlaPheAla 260  
 Db 752 GTGAAGGCCACCGTTTCCCTGCGGAGTCCACGTGTTTCACTTCAGCACCGCTTTGCC 811  
 Qy 261 ArgValAspGluAlaLeuGlyArgProGlyGlyLeuAlaValLeuAlaAlaPheLeuGlu 280  
 Db 812 AGAGTTGACGAGGCTTTGGGGCGCGCGGAGGCTGCGCGGTGTTGGCGCGCTTCTGAG 871

Qy 281 GluGlyProGluGluAsnSerAlaTyrGluGlnLeuLeuSerArgLeuGluGluAla 300  
 Db 872 GAGGCCCCGGAAGAAACACTGCTATGAGCAGTTGCTGCTCGTTTGAAGAATCGCT 931  
 Qy 301 GluGluGlySerGluThrGlnValProGlyLeuAspLysSerAlaLeuLeuProSerAsp 320  
 Db 932 GAGGAAGGCTCAGAGACTCAGGTCCAGGATCCAGGACTGACATATCTGCACCTCTCTGAC 991  
 Qy 321 PheSerArgTyrPheGlnTyrGluGlySerLeuThrProProCysAlaGlnGlyVal 340  
 Db 992 TTGAGCGCTGCTTCCCAATATAGGGGTCTCTGACTACCGCCCTGTGCCAGGGTGTG 1051  
 Qy 341 IleTrpThrValPheAsnGlnThrValMetLeuSerAlaLysGlnLeuHisThrLeuSer 360  
 Db 1052 ATCTGGACTGTGTTTAAACACAGACAGTGTCTGAGTGTCTAAGCAGCTCCACACCTCTCT 1111  
 Qy 361 AspThrLeuTrpGlyProGlyAspSerArgLeuGlnLeuAsnPheArgAlaThrGlnPro 380  
 Db 1112 GACACCTGTGGGACCTGGTGACTCTCGGTACAGCTGAATCTCCAGCGACGACGCT 1171  
 Qy 381 LeuAsnGlyArgValIleGluAlaSerPheProAlaGlyValAspSerSerProArgAla 400  
 Db 1172 TTGAATGGGAGTGAATGAGGCTCTCTCCCTGCTGAGTGTGACAGCTCTCTCGGGCT 1231  
 Qy 401 AlaGluProValGlnLeuAsnSerCysLeuAlaAlaGlyAspLeuAlaLeuValPhe 420  
 Db 1232 GCTGAGCCACTCCAGCTGAATCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1291  
 Qy 421 GlyLeuLeuPheAlaValThrSerValAlaPheLeuValGlnMetArgArgGlnHisArg 440  
 Db 1292 GGCCTCTCTTTTGTGTCCACGCGTCCGCTTCTGTCGAGATGAGAGGCGACGACAGA 1351  
 Qy 441 ArgGlyThrLysGlyValSerTyrArgProAlaGluValAlaGluThrGlyVala 459  
 Db 1352 AGGGAACCAAGAGGGGTGTGAGCTACCGCCAGCAGAGGTAGCCGAGCTGGAGCC 1408

RESULT 3  
 CR597234 1541 bp mRNA linear HTC 21-JUL-2004  
 LOCUS full-length cDNA clone CS0DI085YF18 of Placentia Cot 25-normalized  
 DEFINITION of Homo sapiens (human).  
 CR597234  
 VERSION CR597234.1 GI:50478041  
 KEYWORDS HTC; CNSLT cDNA.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
 Hominidae; Homo.  
 1 (bases 1 to 1541)  
 Li, W.B., Gruber, C., Jessee, J. and Polayes, D.  
 Full-length cDNA libraries and normalization  
 Unpublished  
 Contact : Feng Liang Email : fliang@lifetech.com URL :  
 http://fulllength.invitrogen.com/ Invitrogen Corporation 1600  
 Paraday Avenue  
 2 (bases 1 to 1541)  
 Genoscope.  
 Direct Submission  
 Submitted (20-JUL-2004) Genoscope - Centre National de Sequencage :  
 BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr)  
 - Web : www.genoscope.cns.fr)  
 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime  
 end enriched, double-strand cDNA was digested with Not I and cloned  
 into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library  
 was normalized. Library was constructed by Life technologies, a  
 division of Invitrogen.  
 Location/Qualifiers  
 1..1541  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clones="CS0DI085YF18"

REFERENCE  
 AUTHORS  
 TITLE  
 JOURNAL  
 REMARK

REFERENCE  
 AUTHORS  
 TITLE  
 JOURNAL

COMMENT

FEATURES  
 source

```
/tissue_type="Placenta Cot 25-normalized"  
/plasmid="pCMVSPORT 6"
```

## ORIGIN

**Alignment Scores:**

Pred. No.:	3.12e-195	Length:	1541
Score:	2420.00	Matches:	458
Percent Similarity:	100.0%	Conservative:	1
Best Local Similarity:	99.8%	Mismatches:	0
Query Match:	99.8%	Indels:	0
DB:	4	Gaps:	0

US-09-967-237B-2 (1-459) x CR597234 (1-1541)

Qy	1	MetAlaProLeuGlnCysProSerProTrpLeuProLeuLeuLeuProAlaProGly	20
Db	42	ATGGCTCCCTGTGGCCCAAGCCCCCTGGCTCCCTCTGTGTGATCCCGGGCCCTCGCTCCAGGC	101
Qy	21	LeuThrValGlnLeuLeuLeuSerLeuLeuLeuMetProValHisProGlnArgLeu	40
Db	102	CTCAGCTGTCAACTGCTGCTGTCTCATGTCTTCTGGTGCCCTGTCATCCCCAGAGGTTG	161
Qy	41	ProArgMetGlnGlnAspSerProLeuGlyGlySerSerGlyGluAspAspProLeu	60
Db	162	CCCGGATCAGAGAGATTCCCCCTTGGGAGGAGGCTCTTCTGGGGAGATGACCCACTG	221
Qy	61	GlyGluGluAspLeuProSerGlnGluAspSerProArgGluGluAspProProGlyGlu	80
Db	222	GGCGAGGAGGATCTGCCAGTGAAGAGATTACCACAGAGGAGGATCCACCCGGAGAG	281
Qy	81	GluAspLeuProGlyGluGluAspLeuProGlyGluGluAspLeuProGluValIysPro	100
Db	282	GAGGATCTACCTGGAGAGGAGATCTACCTGGAGAGGAGGATCTACCTGAAGTTTAAAGCCT	341
Qy	101	LysSerGluGluGluGlySerLeuIysLeuGluAspLeuProThrValGluAlaProGly	120
Db	342	AAATCAGAGAGAGGGCTCCCTGAAGTTAGAGGATCTACCTACTGTGTGAGGCTCCCTGGGA	401
Qy	121	AspProGlnGluProGlnAsnAsnAlaHisArgAspLysGluGlyAspAspGlnSerHis	140
Db	402	GATCCTCAGAGACCCTCAGAAATTAATGCCACAGGAGCAAGAAGGGGATGACCCAGAGTCAT	461
Qy	141	TrpArgTyrGlyGlyAspProProTrpProArgValSerProAlaCysAlaGlyArgPhe	160
Db	462	TGGCGCTATGGAGGCGACACCGCCCTGGCCCGCGGTGTCCCCAGAGCTGCGCGGGCCGCTTC	521
Qy	161	GlnSerProValAspIleArgProGlnLeuAlaPheCysProAlaLeuAspGlyProLeu	180
Db	522	CAGTCCCGGGTGATATCCGCCCCCGAGCTCGCCGCCCTTCTCGCGGCCCTTCGCGCCCCCTG	581
Qy	181	GluLeuLeuGlyPheGlnLeuProProLeuProGluLeuArgLeuArgAsnAsnGlyHis	200
Db	582	GAACTCTCTGGCTCTCCAGCTCCCGCGCTCCCGAGAACTGGCGCTCGCAACAATGGCCAC	641
Qy	201	SerValGlnLeuThrLeuProProGlyLysLeuGluMetalaLeuGlyProGlyArgGluTyr	220
Db	642	AGTGTCAACTGACCCCTGCCTCTCGGGCTAGAGATGGCTCTGGGTCCCGGGCGGGAGTAC	701
Qy	221	ArgAlaLeuGlnLeuHisIleuHisTrpGlyAlaAlaGlyArgProGlySerGluHisThr	240
Db	702	CGGGCTCTGGAGCTGCATCTGCATCTGGGGGGCTGCAGGTCGTTCGGGGCTCGAGCACACT	761
Qy	241	ValGluGlyHisArgPheProAlaGluIleHisValValHisLeuSerThrAlaPheAla	260
Db	762	GTGGAAAGGCCACCGTTTCCCTGCCAGATCCACGTGGTTTCACTCAGCACCGCCCTTTTGCC	821
Qy	261	ArgValAspGluAlaLeuGlyArgProGlyGlyLeuAlaValLeuAlaAlaPheLeuGlu	280
Db	822	AGATTGTACAGAGCCCTTGGGGCGCCCGGAGAGCCCTGGCCGTGTGGCCGCTTTCGAG	881
Qy	281	GluGlyProGluGluAsnSerAlaTyrGluGlnLeuLeuLeuSerArgLeuGluGluIleAla	300
Db	882	GAGGGCCCGGAGAAACAGTGCCTATGAGCAGTGTGCTGTCTCGTTTGGAAAGAAATCGCT	941

Qy	301	GluGlySerGlnThrGlnValProGlyLeuAspIleSerAlaLeuLeuProSerAsp	320
Db	942	GAGGAAGGCTCAGAGACTCAGGTCACAGGACTCGACATATCTGCATCTCCCTCGCCCTCTGAC	1001
Qy	321	PheSerArgTyrPheGlnTyrGluGlySerLeuThrThrProProCysAlaGlnGlyVal	340
Db	1002	TTGAGCGCGTACTTCCATATATGAGGGGTCTCTGACTACACCGCCCTGTGCCAGGGGTCTC	1061
Qy	341	IleTrpThrValPheAsnGlnThrValMetLeuSerAlaGlyGlnLeuHisThrLeuSer	360
Db	1062	ATCTGGACTGTGTTTAAACACAGACAGTATCTCTGAGTCTTAAGCAGCTCCACACCCCTCTCT	1121
Qy	361	AspThrLeuTrpGlyProGlyAspSerArgLeuGlnLeuAsnPheArgAlaThrGlnPro	380
Db	1122	GACACCTGTGGGGACCTGTGTGACTCTCGGCTACAGCTGAACCTCCGAGGAGCGAGCCT	1181
Qy	381	LeuAsnGlyArgValIleGluAlaSerPheProAlaGlyValAspSerSerProArgAla	400
Db	1182	TTGATGGGCGAGTGATTGAGGCTCTCTTCCCTGCTGGAGTGGACAGCAGTCTCTCGGCT	1241
Qy	401	AlaGluProValGlnLeuAsnSerCysLeuAlaAlaGlyAspIleLeuAlaLeuValPhe	420
Db	1242	GCTGAGCCAGTCCAGCTGAATTCCTGCTGGCTGCTGCTGACATCTCTAGCCCTGGTCTTT	1301
Qy	421	GlyLeuLeuPheAlaValThrSerValAlaPheLeuValGlnMetArgArgGlnHisArg	440
Db	1302	GGCCTCTCTTTTGTGTGCACGCGTGGCGTCTCTTGTGCAGATGAGAGGCGACGACAGA	1361
Qy	441	ArgGlyThrIysGlyValSerTyrArgProAlaGluValAlaGluThrGlyVala	459
Db	1362	AGGGGAACCAAGGGGGTGTGAGTACCGCCAGCAGAGAGGTAGCCGAGAGCTGGAGCC	1418
RESULT 4	CR590646	1492 bp mRNA linear HTC 21-JUL-2004	
LOCUS	full-length cDNA clone CS0DJ005YK14 of T cells (Jurkat cell line)		
DEFINITION	Cot 10-normalized of Homo sapiens (human).		
ACCESSION	CR590646		
VERSION	CR590646.1	GI:50471453	
KEYWORDS	HTC; CNSLT_cDNA.		
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
AUTHORS	Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;		
JOURNAL	Hominidae; Homo.		
REMARK	1 (bases 1 to 1492)		
	Li, W.B., Gruber, C., Jessee, J. and Polayes, D.		
	Full-length cDNA libraries and normalization		
	Unpublished		
	Contact: Feng Liang Email: fliang@lifetech.com URL:		
	http://fulllength.invitrogen.com/ invitrogen Corporation 1600		
	Faraday Avenue		
	2 (bases 1 to 1492)		
REFERENCE	Genoscope.		
AUTHORS	Direct Submission		
TITLE	Submitted (20-JUL-2004) Genoscope - Centre National de Sequences:		
JOURNAL	BP 191 91006 EVRY cedex - FRANCE (E-mail: seqref@genoscope.cns.fr)		
COMMENT	- Web: www.genoscope.cns.fr)		
	1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime		
	end enriched, double-strand cDNA was digested with Not I and cloned		
	into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library		
	was normalized. Library was constructed by Life Technologies, a		
	division of invitrogen.		
FEATURES	Location/Qualifiers		
source	1. .1492		
	/organism="Homo sapiens"		
	/mol_type="mRNA"		
	/db_xref="taxon:9606"		
	/clone="CS0DJ005YK14"		
	/tissue_type="T cells (Jurkat cell line) Cot		
	10-normalized"		
	/plasmid="pCMVSPORT 6"		

## ORIGIN

## Alignment Scores:

Pred. No.: 1.63e-180 Length: 1492  
Score: 2361.50 Matches: 449  
Percent Similarity: 98.0% Conservatives: 1  
Best Local Similarity: 97.8% Mismatches: 0  
Query Match: 97.4% Indels: 9  
DB: 4 Gaps: 1

US-09-967-237B-2 (1-459) x CR590646 (1-1492)

QY 1 MetAlaProLeuCySProSerProTriPLeuProLeuLeuLeuLeuProAlaProGly 20  
DB 43 ATGGCTCCCTGTGCCCCAGCCCTGGCTCCCTCTGTTGATCCGGCCCTGTCTCAGGC 102  
QY 21 LeuThrValGlnLeuLeuLeuSerLeuLeuLeuMetProValHisProGlnArgLeu 40  
DB 103 CTCACTGTCAACTGCTGCTGCTCACTGCTGCTTCTGCTGCTGCTTCCATCCCCAGAGGTTG 162  
QY 41 ProArgMetGlnGluAspSerProLeuGlyGlyGlySerSerGlyGluAspProLeu 60  
DB 163 CCCCGAGTGCAGAGGATTCCTCCCTGGAGGAGGCTCTTCTGGGGAAGATGACCCACTG 222  
QY 61 GlyGluGluAspLeuProSerGluGluAspSerProArgGluGluAspProGlyGlu 80  
DB 223 GCGGAGGAGATCTGCCAGTGAAGAGATTCACCCAGAGAGGAGGATTCACCCGGAGAG 282  
QY 81 GluAspLeuProGlyGluGluAspLeuProGlyGluGluAspLeuProGluValLysPro 100  
DB 283 GAGGATCTACTCGAGAGGAGGATCTACTCGGAGGAGGATCTACTGAAAGTTAGCCT 342  
QY 101 LysSerGluGluGlySerLeuLysLeuGluAspLeuProThrValGluAlaProGly 120  
DB 343 AAATCAGAAGAAGAGGCTCCCTGAAGTTAGAGATCTACTACTGTTGAGGCTCTCGGA 402  
QY 121 AspProGlnGluProGlnAsnAlaHisArgAspLysGluGlyAspAspGlnSerHis 140  
DB 403 GATCTCAGAACCCCAAGATATATGCCACAGGACAAAGAGGGGATGACAGATCAT 462  
QY 141 TrpArgTyrGlyAspProProTriPProArgValSerProAlaCySAlaGlyArgPhe 160  
DB 463 TGGCGCTATGAGGCGACCCGCTCGCCCGGGGTGCTCCAGCGCTCGCGGGCGGCTTC 522  
QY 161 GlnSerProValAspLeuArgProGlnLeuAlaPheCysProAlaLeuArgProLeu 180  
DB 523 CAGTCCCGGTGGATATCCGCCCGCCAGCTCGCGCCCTTCTGCCGGGCGCTCGCGCCCTG 582  
QY 181 GluLeuLeuGlyPheGlnLeuProProLeuProGluLeuArgLeuAsnGlyHis 200  
DB 583 GAATCTCTGGGCTTCCAGCTCCCGCCGCTCCAGACTGGCGCTGCCAACAATGGCCAC 642  
QY 201 SerValGlnLeuThrLeuProProGlyLeuGluMetAlaLeuGlyProGlyArgGluTyr 220  
DB 643 AGTGTGCAACTGACCTGCTCTCGGGCTAGAGATGCTCTGGGTCCCGGGCGGAGTAC 702  
QY 221 ArgAlaLeuGlnLeuHisLeuHisTyrGlyAlaAlaGlyArgProGlySerGluHisThr 240  
DB 703 CGGCTCTGAGCTGCACTGCACTGGGGGGCTGAGGTGCTCGGGCTCGAGACACT 762  
QY 241 ValGluGlyHisArgPheProAlaGluHisValValHisLeuSerThrAlaPheAla 260  
DB 763 GTGGAGGCCACCGTTTCCCTGCCAGATCCAGTGGTTTACCTCAGCAGCGCTTTGCC 822  
QY 261 ArgValAspGluAlaLeuGlyArgProGlyGlyLeuAlaValLeuAlaPheLeuGlu 280  
DB 823 AGAGTTGACGAGGCTTTGGGGCGCCCGGAGGCTGCGCGTGTGGCGCTTCTCGAG 882  
QY 281 GluGlyProGluGluAsnSerAlaTyrGluGlnLeuSerArgLeuGluGluAla 300  
DB 883 GAGGCGCCGGAAGAAAACAGTGCCTATGAGCAGTGTGCTGCTCGCTTTGGAAGAAATCGCT 942  
QY 301 GluGluGlySerGluThrGlnValProGlyLeuAspPheSerAlaLeuLeuProSerAsp 320

DB 943 GAGGAAGGCTCAGAGACTCAGGTCCCGAGACTGACATATCTGCACCTCTCCCTCTGAC 1002  
QY 321 PheSerArgTyrPheGlnTyrGluGlySerLeuThrProProCysAlaGlnGlyVal 340  
DB 1003 TTCAAGCGCTACTTCCAAATATGAGGGGTCTCTGACTACACCGCCCTGTGCCCCAGGGTGC 1062  
QY 341 IleTrpThrValPheAsnGlnThrValMetLeuSerAlaLysGlnLeuHisThrLeuSer 360  
DB 1063 ATCTGAGCTGTGTTTAAACACAGACAGTGTGTGCTGTGCTTANGCAGCTCCACACCTCTCT 1122  
QY 361 AspThrLeuTrpGlyProGlyAspSerArgLeuGlnLeuAsnPheArgAlaThrGlnPro 380  
DB 1123 GACACCTGTGGGACCTCTGGTACTCTCGCTACAGCTGAACCTCCGAGCGAGCAGCCT 1182  
QY 381 LeuAsnGlyArgValIleGluAlaSerPheProAlaGlyValAspSerProArgAla 400  
DB 1183 TTGAATGGGCGAGTGAATTGAGGCTCTCTCCCTGCTGGAGTGGACAGCAGTCTCTCGGGCT 1242  
QY 401 AlaGluProValGlnLeuAsnSerCysLeuAlaAlaGlyAspIleLeuAlaLeuValPhe 420  
DB 1243 GCTGAGCCA-----GGTGACATCCTAGCCCTGGTTTTT 1275  
QY 421 GlyLeuLeuPheAlaValThrSerValAlaPheLeuValGlnMetArgArgGlnHisArg 440  
DB 1276 GGCTCTCTTTTCTGCTCACCAGCTCGCTTCTTGTGCAGATGAGAGGCGAGCAGAGA 1335  
QY 441 ArgGlyThrLysGlyGlyValSerTyrArgProAlaGluValAlaGluThrGlyAla 459  
DB 1336 AGGGGAACCAAGGGGTGTGAGCTACCGCCAGCAGAGGTAGCCGAGACTGGAGCC 1392  
RESULT 5  
DQ049266 1380 bp DNA linear GSS 02-JUN-2005  
LOCUS Homo sapiens CA9 gene, VIRTUAL TRANSCRIPT, partial sequence,  
DEFINITION genomic survey sequence.  
ACCESSION DQ049266  
VERSION DQ049266.1 GI:66902465  
KEYWORDS GSS.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
REFERENCE 1 (bases 1 to 1380)  
AUTHORS Nielsen,R., Bustamante,C., Clark,A.G., Glanowski,S., Sackton,T.B.,  
Hubisz,M.J., Fiedel-Alon,A., Tanenbaum,D.M., Civeillo,D.,  
White,T.J., Sninsky,J.J., Adams,M.D. and Cargill,M.  
TITLE A Scan for Positively Selected Genes in the Genomes of Humans and  
Chimpanzees  
JOURNAL (er) PLoS Biol. 3 (6), E170 (2005)  
PUBMED 15869325  
REFERENCE 2 (bases 1 to 1380)  
AUTHORS Nielsen,R., Bustamante,C., Clark,A.G., Glanowski,S., Sackton,T.B.,  
Hubisz,M.J., Fiedel-Alon,A., Tanenbaum,D.M., Civeillo,D.,  
White,T.J., Sninsky,J.J., Adams,M.D. and Cargill,M.  
TITLE Direct Submission  
JOURNAL Submitted (05-MAY-2005) Celera Genomics, 45 West Gude Drive,  
Rockville, MD 20850, USA  
COMMENT This sequence was made by sequencing genomic exons and ordering  
them based on alignment. Translation starts at the beginning of  
alignment.  
FEATURES  
source Location/Qualifiers  
1..1380  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
/chromosome="9"  
c1..>1380  
/gene="CA9"  
/locus\_tag="HC17258"  
gene  
ORIGIN

Alignment Scores:			
Pred. No.:	1.27e-165	Length:	1380
Score:	2177.00	Matches:	422
Percent Similarity:	92.2%	Conservative:	1
Best Local Similarity:	91.9%	Mismatches:	36
Query Match:	89.8%	Indels:	0
DB:	11	Gaps:	0
US-09-967-237B-2 (1-459) x DQ049266 (1-1380)			
QY	1	MetAlaProLeuCysProSerProTTPLeuProLeuLeuLeuProAlaProAlaProGly	20
Db	1	ATGGCTCCCTGTGGCCAGCCCTGGCTCCCTCTGTGTGATCCCGGCCCTGTCTCAGGC	60
QY	21	LeuThrValGlnLeuLeuSerLeuLeuLeuMetProValHisProGlnArgLeu	40
Db	61	CTCACTGTGCAACTGCTGCTGTCTACTGCTCTCTGGTGCTGTCCATCCCCAGAGTTG	120
QY	41	ProArgMetGlnGluAspSerProLeuGlyGlyGlySerSerGlyGluAspAspProLeu	60
Db	121	CCCCGGATGCAGGAGATTCCCTCTGGGAGGAGGCTCTTCTGGGGAAGATGACCCACTG	180
QY	61	GlyGluGluAspLeuProSerGluGluAspSerProArgGluGluAspProProGlyGlu	80
Db	181	GGCGAGAGGATCTGCCAGTGAAGAGGATTCCACCAGAGAGGAGGATCCACCCGAGAG	240
QY	81	GluAspLeuProGlyGluGluAspLeuProGlyGluGluAspLeuProGluValHisPro	100
Db	241	GAGGATCACTGGAGGAGGAGTCTACCTGGAGAGGAGGATCTACTGAAGTTAAGCCT	300
QY	101	LySerGluGluGlySerLeuLysLeuGluAspLeuProThrValGluAlaProGly	120
Db	301	AAATCAGAGAGGCGTCCCTGAAAGTTAGAGGATCTACTACTGTGAGGCTCTCTGGA	360
QY	121	AspProGlnGluProGlnAenAenAlaHisArgAspLysGlyGlyAspAspGlnSerHis	140
Db	361	GATCTCTCAAGAACCCAGAAATATGCCACAGGACAAGAGGGGATGACCAAGATCAT	420
QY	141	TrpArgTyrGlyGlyAspProProTTPProArgValSerProAlaCysAlaGlyArgPhe	160
Db	421	TGGCGCTATGGAGNN	480
QY	161	GlnSerProValAspLeArgProGlnLeuAlaAlaPheCysProAlaLeuArgProLeu	180
Db	481	NNNG	540
QY	181	GluLeuLeuGlyPheGlnLeuProProLeuProGluLeuArgLeuArgAenAnglyHis	200
Db	541	GAACTCTGGGCTTCCAGCTCCCGCGCTCCAGAACTGGCGCTGCCAACAATGGCCAC	600
QY	201	SerValGlnLeuThrLeuProProGlyLeuGluMetAlaLeuGlyProGlyArgGluTyr	220
Db	601	AGTGTGCAACTGACCTGCTCTGGCTAGAGATGGCTCTGGGTCCCGGCGGGAGTAC	660
QY	221	ArgAlaLeuGlnLeuHisLeuHisTrpGlyAlaAlaGlyArgProGlySerGluHisThr	240
Db	661	CGGGCTCTGCAGCTGCATCTGCATCGGGGGCTGCAGGTCTCGCGGCTCGAGCACACT	720
QY	241	ValGluGlyHisArgPheProAlaGluLeHisValValHisLeuSerThrAlaPheAla	260
Db	721	GTGGAAGGCCACCGTTTCCCTGCGGAGATCCAGTGGTTCACTCAGCACCGGCTTTGCC	780
QY	261	ArgValAspGluAlaLeuGlyArgProGlyGlyLeuAlaValLeuAlaAlaPheLeuGlu	280
Db	781	AGAGTTGACGAGGCTTTGGGGGCCCGGGAGGCTGCGCGTGTGGCGCTTCTCGAG	840
QY	281	GluGlyProGluGlnSerAlaTyrGluGlnLeuLeuSerArgLeuGluGluLeuAla	300
Db	841	GAGGGCCCGAAGAAACAGTGCCTATGACAGTGTCTGTCTGCTTGGGAAGAAATCGCT	900
QY	301	GluGluGlySerGluThrGlnValProGlyLeuAspLysSerAlaLeuLeuProSerAsp	320
Db	901	GAGGAAGGCTCAGAGACTCAGGTCCAGGACTGGACATATCTGCATCTCTGCCCTCTGAC	960
RESULT 6			
DQ049267			
LOCUS			
DEFINITION			
ACCESSION			
VERSION			
KEYWORDS			
SOURCE			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
COMMENT			
FEATURES			
source			
gene			
ORIGIN			
Alignment Scores:			
Pred. No.:			
Score:			
Length:			
1380			
Matches:			
1926.00			
Length:			
1380			
Matches:			
385			



```
Score: 1790.00 Matches: 337
Percent Similarity: 98.0% Conservative: 4
Best Local Similarity: 96.8% Mismatches: 6
Query Match: 73.8% Indels: 1
DB: 1 Gaps: 0

US-09-967-237B-2 (1-459) x AL542336 (1-1084)

QY 3 ProLeuCysProSerProTTPLeuProLeuLeuLeuLeuProAlaProAlaProGlyLeuThr 22
DB 29 CCCTGTGCCCCAGCCCTCGCTCCCTCTGTGTATCCCGGCCCTGTCTCAGGCTCACT 88
QY 23 ValGlnLeuLeuSerLeuLeuLeuLeuLeuMetProValHisProGlnArgLeuProArg 42
DB 89 GTGCAACTGCTGTCTCACTGCTGCTTCTGTGCTGCTCTCATCCACAGAGTTGCCCGG 148
QY 43 MetGlnLeuAspSerProLeuGlyGlySerSerGlyGluAspAspProLeuGlyGlu 62
DB 149 ATGAGAGGAGATTCCCTTGGGAGGAGGCTCTTCTGGGGAAGATGATGATCCCACTGGGCGAG 208
QY 63 GluAspLeuProSerGluGluAspSerProArgGluGluAspProProGlyGluGluAsp 82
DB 209 GAGGATCTGCCAGTGAAGAGGATTACCCAGAGAGGAGGATCCACCCGAGAGGAGGAT 268
QY 83 LeuProGlyGluGluAspLeuProGlyGluGluAspLeuProGluValLysProLysSer 102
DB 269 CTACCTGGAGAGGAGGATCTACCTGGAGAGGAGGATCTACCTGAAGTTAAGCCTAAATCA 328
QY 103 GluGluGluGlySerLeuLysLeuGluAspLeuProThrValGluAlaProGlyAspPro 122
DB 329 GAAGAAGAGGCTCCCTGGAAGTTAGAGGATCTACCTACTGTGAGGCTCTTGAGATCCT 388
QY 123 GlnGluProGlnAenAenAlaHisArgAspLysGluGlyAspAspGlnSerHisTrpArg 142
DB 389 CAGNACCCAGATATATGCCACAGGGACAAAGAGGGGATGACACAGATCATTTGGCGC 448
QY 143 TyrGlyGlyAspProProTTPProArgValSerProAlaCysAlaGlyArgPheGlnSer 162
DB 449 TATGAGGCGCAGCCGCTGCGCCCGGCTGTCGCCAGCTGCGCGGCGCTTCCAGTCC 508
QY 163 ProValAspLeuArgProGlnLeuAlaAlaPheCysProAlaLeuArgProLeuGluLeu 182
DB 509 CCGGTGGATATCCGCCCCAGCTGCGCCCTTCTGCGCCGCGCTGCGCCCGCTGGAATC 568
QY 183 LeuGlyPheGlnLeuProProLeuProGluLeuArgLeuArgAenAenGlyHisSerVal 202
DB 569 CTGGGCTTCCAGTCTCCCGCGCTCCCAAGACTGGCGCTGGCAACATGSCCAGATGTG 628
QY 203 GlnLeuThrLeuProProGlyLeuGluMetAlaLeuGlyProGlyArgGluTyrArgAla 222
DB 629 CAATGACCTGCTCTGGGCTAGAGATGGCTCTGGGCTCCCGGCGGAGTACCGGGCT 688
QY 223 LeuGlnLeuHisLeuHisTrpGlyAlaAlaGlyArgProGlySerGluHisThrValGlu 242
DB 689 CTGAGCTGATCTGCACTGGGGGCTGCGAGGCTGTCGGGCTCGGAGCACTGTGGAA 748
QY 243 GlyHisArgPheProAlaGluLeuHisValValHisLeuSerThrAlaPheAlaArgVal 262
DB 749 GGCTACCTGTTCCCTGCGGAGATCCACGTGTTCACTCAGACCGGCTTTGGCAGAGTT 808
QY 263 AspGluAlaLeuGlyArgProGlyGlyLeuAlaValLeuAlaPheLeuGluGluGly 282
DB 809 GACGAGGCTTGGGGCGCCGGAGGCTGCGCGTGTGTGGCGCTTCTTGAGAGAGGC 868
QY 283 ProGluGluAenSerAlaTyrGluGlnLeuSerArgLeuGluLeuLeuAlaGluGlu 302
DB 869 CCGGAAGAAACAGTGCTATGAGCAGTTGCTGTCTGCTTGGGAAGAAATCGCTGAGAA 928
QY 303 GlySerGluThrGlnValProGly-LeuAspLeuSerAlaLeuLeuProSerAspPheSe 322
DB 929 GGCTCAGAGACTCAGGTCACGAGACTTGGACATATCTKSACTCTGSCCTCTGACTTCAG 988
QY 322 rArgTyrPheGlnTyrGluGlySerLeuThrThrProProCysAlaGlnGlyValIleTr 342
```

```
Db 989 CCGCTACTTCAAAATATGAGGGGCTCTGACTACACGCCCTGTGCGAGGGTGTCTCATCTK 1048
QY 342 pThrValPheAsnGlnThrVal 349
Db 1049 GACTGTGTTTAAACAGAMAGTK 1070

RESULT 8
AL554705 1070 bp mRNA linear EST 30-MAR-2004
LOCUS AL554705 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA
DEFINITION clone CS001085Yf18 5-PRIME, mRNA sequence.
ACCESSION AL554705
VERSION AL554705.3 GI:45859455
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homnidae; Homo.
REFERENCE 1 (bases 1 to 1070)
AUTHORS Li, W.B., Gruber, C., Jesse, J. and Polayes, D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)
COMMENT On Feb 15, 2001 this sequence version replaced gi:31276515.
Contact: Genoscope
Genoscope - Centre National de Sequencage
2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
end enriched, double-strand cDNA was digested with Not I and cloned
into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
was normalized. Library was constructed by Life Technologies, a
division of Invitrogen. This sequence belongs to sequence cluster
5300.f
For more information about this cluster, see
http://www.genoscope.cns.fr/cdna?s=CS001085DC090P1&e=5300.f.
FEATURES
Location/Qualifiers
source
1..1070
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS001085Yf18"
/tissue_type="PLACENTA COT 25-NORMALIZED"
/notes="1st strand cDNA was primed with a NotI-oligo(dT)
primer. Five prime end enriched, double-strand cDNA was
digested with Not I and cloned into the Not I and EcoR V
sites of the pCMVSPORT 6 vector. Library was normalized."

Alignment Scores:
Pred. No.: 2-2e-116 Length: 1070
Score: 1567.00 Matches: 318
Percent Similarity: 92.5% Conservative: 4
Best Local Similarity: 91.4% Mismatches: 22
Query Match: 64.6% Indels: 8
DB: 1 Gaps: 2

US-09-967-237B-2 (1-459) x AL554705 (1-1070)

QY 1 MetAlaProLeuCysProSerProTTPLeuProLeuLeuLeuProAlaProAlaProGly 20
DB 42 ATGGCTCCCCGTGCGCCAGCCCTCGCTCCCTCTGTGTATCCCGGCCCTGTCTCCAGGC 101
QY 21 LeuThrValGlnLeuLeuLeuSerLeuLeuLeuMetProValHisProGlnArgLeu 40
DB 102 CTCACTGTGCAACTGCTGTCTCACTGCTCTTCTGTGCTCTGTCCATCCCAAGAGTTG 161
QY 41 ProArgMetGlnGluAspSerProLeuGlyGlyGlySerSerGlyGluAspAspProLeu 60
DB 162 CCGCGATGACAGAGGATTCCTCCCTTGGGAGGAGGCTCTTCTGGGAAGATGATGCCACTG 221
```



Db	463	TGGCGCTATGAGAGCGACCCCGCCTTGGCGCCCGGGTGTCCTCCACGCTCTGGCGGGCGCGCTTC	522	
Qy	161	GlnSerProValAspIleAArgProGlnLeuAlaAaPheCysProAlaLeuArgProLeu	180	
Db	523	CAGTCCCGGTGGATATCCGCCCCCAGCTCGCGCCTTCTGCGCGGCGCTGCGCCCCCTG	582	
Qy	181	GluLeuLeuGlyPheGlnLeuProProLeuProGlnLeuLeuArgLeuArgAenAenGlyHis	200	
Db	583	GAACCTCTGGGCTTCAGCTCCCGCGCTCCAGAACTGGCGCTTGGCGCAACATGGCCAC	642	
Qy	201	SerValGlnLeuThrLeuProProGlyLeuGluMetAlaLeuGlyProGlyArgGluTyr	220	
Db	643	AGTGTGCAACTGACCTCGCTCTGGCTAGAGATGGCTCTGGGTCCCGGGCGGAGTAC	702	
Qy	221	ArgAlaLeuGlnLeuHisLeuHisTirpGlyAlaAlaGlyArgProGlySerGluHisThr	240	
Db	703	CGGGCTCTGCAGCTGATCTGCATCTGGGGGGCTGCAGGTGCTCGGGCTCGAGACACT	762	
Qy	241	ValGluGlyHisArgPheProAlaGluLeuHisValValHisLeuSerThrAlaPheAla	260	
Db	763	GTGGAAGGCCACCGTTTCCCTGCCGAGATCCACGTGTTTCACTTCAGCACCGCTTTGCC	822	
Qy	261	ArgValAspGluAlaLeuGlyArgProGlyGlyLeuAlaValLeuAlaAlaPheLeuGlu	280	
Db	823	AGAGTTTGACAGGCGCTTGGGGCGCCGGAGGCGCTGGCGGTGTGGC-GCCTTTCTGGAG	881	
Qy	281	GluGlyProGluGluAsnSerAlaTyrGlnGlnLeuLeuSerArgLeuGluGluLeuAla	300	
Db	882	GAGGGSCCGGAGA-AACAGTGCTTATGAGCAGTGTCTGTCTGCTTGGHARA-ATCGCT	939	
Qy	301	GluGluGlySerGluThrGlnValProGlyLeuAspIleSerAlaLeuLeuPro	318	
Db	940	GAGGAAGC-TCAGAGACTCAGKCCAGGACTGGACAWATT-GCACTCTCTGCT	991	
RESULT 10	EX401186	1017 bp	linear	EST 29-APR-2000
LOCUS	EX401186	Homo sapiens HELA CELLS	COT 25-NORMALIZED	Homo sapiens
DEFINITION	CDNA clone CS0DK009YP10 5-PRIME, mRNA	sequence.		
ACCESSION	EX401186			
VERSION	EX401186.2	GI:46876709		
KEYWORDS	EST.			
SOURCE	Homo sapiens			
ORGANISM	Homo sapiens (human)			
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
	Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;			
	Homnidae; Homo.			
REFERENCE	1 (bases 1 to 1017)			
AUTHORS	Li, W.B., Gruber, C., Jessee, J. and Polayes, D.			
TITLE	Full-length cDNA libraries and normalization			
JOURNAL	Unpublished (2001)			
COMMENT	On May 13, 2003 this sequence version replaced gi:30626393.			
	Contact: Genoscope			
	Genoscope - Centre National de Sequencage			
	2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE			
	Email: seqrefgenoscope.cns.fr, Web : www.genoscope.cns.fr			
	1st strand cDNA was primed with a NotI-oligo(dT) primer. Five primer			
	into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library			
	was normalized. Library was constructed by Life Technologies, a			
	division of Invitrogen. This sequence belongs to sequence cluster			
	5300.f			
	For more information about this cluster, see			
	http://www.genoscope.cns.fr/cdna?c=CS0DK009DH05QP1&c=5300.f.			
FEATURES	Location/Qualifiers			
source	1..1017			
	/organism="Homo sapiens"			
	/mol_type="mRNA"			
	/db_xref="taxon:9606"			
	/clone="CS0DK009YP10"			
	/cell_type="HELA CELLS			
	/cell_line="HELA"			
	/clone_lib="Homo sapiens HELA CELLS			
	COT 25-NORMALIZED"			

```

855 CCCGAGAAACAGTGCCTATGACAG--TTGVTCTCGCTGGWAGAAATCGCTGAGGAA 911
|||||
303 GlycerGluThrGlnValProGlyLeuAspIleSerAlaLeu 316
|||||
912 GGT---CAGAGMTCAAGTCCAGGACTGGAMTAYTGCVTYCTG 950
|||||

RESULT 11
BX423970      916 bp      mRNA      linear      EST 03-MAY-2004
LOCUS        BX423970 Homo sapiens NEUROBLASTOMA Homo sapiens cDNA clone
DEFINITION   CS0DA003YB12 5-PRIME, mRNA sequence.
ACCESSION    BX423970
VERSION      BX423970.2 GI:46955310
KEYWORDS     EST.
SOURCE       Homo sapiens (human)
ORGANISM     Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
1 (bases 1 to 916)
Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
Full-length cDNA libraries and normalization
Unpublished (2001)
On May 15, 2003 this sequence version replaced gi:30766328.
Contact: Genoscope
Genoscope - Centre National de Sequencage
2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
Email: segre@genoscope.cns.fr, Web : www.genoscope.cns.fr
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
end enriched, double-strand cDNA was digested with Not I and cloned
into the Not I and EcoRV sites of the pCMVSPORT 6 vector. Library
was not normalized. Library was constructed by Life Technologies, a
division of Invitrogen.
This sequence belongs to sequence cluster 5300.f
For more information about this cluster, see
http://www.genoscope.cns.fr/cdna?s=CS1DA001ZF09QP1&c=5300.f.

FEATURES
Location/Qualifiers
1..916
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DA003YB12"
/tissue_type="NEUROBLASTOMA"
/clone_lib="Homo sapiens NEUROBLASTOMA"
/notes="Vector: pCMVSPORT 6; 1st strand cDNA was primed
with a NotI-oligo(dT) primer. Five prime end enriched,
double-strand cDNA was digested with Not I and cloned into
the Not I and EcoRV sites of the pCMVSPORT 6 vector.
Library was not normalized."

ORIGIN
Alignment Scores:
Pred. No.:      8.75e-109      Length:      916
Score:          1472.00      Matches:      276
Percent Similarity: 92.4%      Conservative: 5
Best Local Similarity: 90.8%      Mismatches: 23
Query Match:    60.7%      Indels:      0
Gaps:          5

US-09-967-237B-2 (1-459) x BX423970 (1-916)

Qy 1 MetlaProLeuCysProSerProTrrProLeuLeuLeuProAlaProAlaProGly 20
Db 5 ATGGCTCCCTGTGCCAGCCCTCGCTCCCTCTGTGTGAWCCCGCCCTCGCTCCAGGC 64
Qy 21 LeuThrValGlnLeuLeuSerLeuLeuLeuMetProValHisProGlnArgLeu 40
Db 65 CTCAGTGTGCACACGCTGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 124
Qy 41 ProArgMetGlnGluAspSerProLeuGlyGlySerSerGlyGluAspProLeu 60
Db 125 CCCCGGATGCAGGAGGATTCCTCCCTTGGGAGAGGCTCTWCTGGGGAAGATGACCCACTG 184

```

```

Qy 61 GlyGluGluAspLeuProSerGluGluAspSerProArgGluGluAspProProGlyGlu 80
Db 185 GCGAGAGAGATCTGCCAGTGAAGAGATTCACCCAGAGAGAGAGATTCACCCGGAGAG 244
Qy 81 GluAspLeuProGlyGluGluAspLeuProGlyGluGluAspLeuProGluValLeuPro 100
Db 245 GAGGATCTACCTGGAGAGGAGGATCTACCTGGAGAGGAGGATCWACCTGAAGTAAAGCCT 304
Qy 101 LysSerGluGluGluGlySerLeuLeuLeuAspLeuProThrValAlaAlaProGly 120
Db 305 AAATCAGAAGAGAGAGGCTCCCTGAAGTGAAGAGATCTACTACTGTGAGAGGCTCTCGGA 364
Qy 121 AspProGlnGluProGlnAsnAlaHisArgAspLysGluGlyAspAspGlnSerHis 140
Db 365 GAACCTCAGAACCCAGATTAAGCCACAGGACCAAGAGAGGAGATGACAGAGTCAT 424
Qy 141 TrpArgTyroGlyAspProProTrrProArgValSerProAlaCysAlaGlyArgPhe 160
Db 425 TGGCGGATGGAGGCGACCCGCTGGCCCGGGGTGTACCCAGCCTCGCGGGCGCGCTTC 484
Qy 161 GlnSerProValAspIleArgProGlnLeuAlaAlaPheCysProAlaLeuArgProLeu 180
Db 485 CAGTCCCGGTGGATATCCGCCGCCAGCTCGCCGCTTCTGCGCGGCTTGGCGCCCTG 544
Qy 181 GluLeuLeuGlyPheGlnLeuProProLeuProGluLeuArgLeuArgAsnGlyHis 200
Db 545 GAACCTCTGGGCTTCAGCTCCCGCGCTCCAGAGATCGCGCTCGCMACNATGGCCAC 604
Qy 201 ServGlnLeuThrLeuProProGlyLeuGluMetAlaLeuGlyProGlyArgGluTyr 220
Db 605 AGTGTGCAACTGACCTGCTCTGGCTAGAGAGGCTCTGGCTCCCGGGAGGTAC 664
Qy 221 ArgAlaLeuGlnLeuHisLeuHisTrrGlyAlaAlaGlyArgProGlySerGluHisThr 240
Db 665 CGGGCTCTGAGCTGCAATCTGCACTGGGGGCTGCAGGCTGCTCGGGCTCGAGCACACT 724
Qy 241 ValGluGlyHisArgPheProAlaGluLeuHisValHisLeuSerThrAlaPheAla 260
Db 725 GTGGARGCCACCGTTTCCCTCGGAGATCCAGTGTGTTTAMCTCACCACCGCTTTGCC 784
Qy 261 ArgValAspGluAlaLeuGlyArgProGlyGlyLeuAlaValLeuAlaAlaPheLeuGlu 280
Db 785 ARAATTGAGGAGGCTTTGGGGCGCCGGGAGGCTGGCCCTGTGGCGGCTTTCTGGAG 844
Qy 281 GluGlyProGluGlnSerAlaTyrGluGlnLeuLeuSerArgLeuGluGluAla 300
Db 845 GAGGGCGCGAARAAAATAATRCCTATRCAGTGTCTTTCGCTTGGRAAAAWMGCT 904
Qy 301 GluGluGlySer 304
Db 905 GAGAGAGGCTCA 916

RESULT 12
COS79387      971 bp      mRNA      linear      EST 20-JUL-2004
LOCUS        ILLUMIGEN MCQ 50239 Katze_WML Macaca mulatta cDNA clone
DEFINITION   IBIUM:17608 5' similar to Bases 6 to 971 highly similar to human
CA9 (Hs.63287), mRNA sequence.
ACCESSION    COS79387
VERSION      COS79387.1 GI:50410317
KEYWORDS     EST.
SOURCE       Macaca mulatta (rhesus monkey)
ORGANISM     Macaca mulatta
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Cercopithecoidea; Cercopithecoidea; Macaca.
1 (bases 1 to 971)
Magnes,C.L., Fellin,P.C., Thomas,M.J., Korth,M.J., Agv,M.B.,
Proli,S.C., Fitzgibbon,M., Scherer,C.A., Miner,D.G., Katze,M.G. and
Iadonato,S.P.
Analysis of the Macaca mulatta transcriptome and the sequence
divergence between Macaca and human

```

JOURNAL Genome Biol. 6 (7), R60 (2005)  
 PUBMED 15998449  
 COMMENT Contact: C. Magnes  
 Illumigen Biosciences Inc.  
 2203 Airport Way S, Suite 450, Seattle, WA 98134, USA  
 Tel: 2063780400  
 Fax: 2063780408  
 Email: cmagnes@illumigen.com  
 Sequenced on 2004.07.15. 725 Q20 bases. Library Preparation: Prof.  
 Michael Katze Lab at University of Washington DNA Sequencing:  
 Illumigen Biosciences Inc. For further information, see  
 http://www.macaque.org

PCR Primers  
 FORWARD: CCTCTACTAAGGGAACAAAA  
 BACKWARD: CACTATAGGGCGAATTGGGTA  
 Insert Length: 971 Std Error: 0.00  
 Plate: CL000413 row: A column: 10  
 Seq primer: CCTCTACTAAGGGAACAAAA  
 POLYA-No.

## FEATURES

source

Location/Qualifiers

1. .971  
 /organism="Macaca mulatta"  
 /mol\_type="mRNA"  
 /strain="Indian"  
 /db\_xref="taxon:9544"  
 /clone="IBIUM:17608"  
 /sex="male"  
 /dev\_stage="adult"  
 /lab\_host="Electromax DH10B"  
 /clone\_lib="Katze MWIL"  
 /note="Organ: ileum; Vector: pDONR 222; Site 1: Berg I;  
 Site 2: Berg I; Created from CloneMiner cDNA Library  
 Construction kit (catalog #18249-029)"

## ORIGIN

Alignment Scores:

Pred. No.: 3.6e-107 Length: 971  
 Score: 1452.50 Matches: 286  
 Percent Similarity: 88.9% Conservative: 11  
 Best Local Similarity: 85.6% Mismatches: 20  
 Query Match: 59.9% Indels: 17  
 DB: 7 Gaps: 2

US-09-967-237B-2 (1-459) x COS79387 (1-971)

QY 55 GlyAlaAspProLeuGlyGluAspLeuProSerGluGluAspSerProArgGlu 74  
 DB 3 GGGAGATGACCCACTGCATGAGGAGGATCTGCCAGTGAAGAGGATCCACCCAGAGAG 62  
 QY 75 GluAspProGlyGluGluAspLeuProGlyGluAspLeuProGlyGluGluAsp 94  
 DB 63 GAGGATCCA-----CCAGAGAGGAGGAT 86  
 QY 95 LeuProGluValIysProIysSerGluGluGlySerLeuIysLeuGluAspLeuPro 114  
 DB 87 CTACCTGAAGTTAAGCTTAATCAGAGAGAGGGCTCCCTGAAGTTAGAGGATCTACCT 146  
 QY 115 ThrValGluAlaProGlyAspProGlnGluProGlnAsnAsnAlaHisArgAspLysGlu 134  
 DB 147 ACTGTTGAGGCTCTCGAGATCCTCAAGAACCCAGAGTAATGCCACAGGAGCAAGAA 206  
 QY 135 GlyAspAspGlnSerHisTrpArgTyrGlyAspProTrpProArgValSerPro 154  
 DB 207 GGGGTAACACAGACACTGGCGCTATGGAGGGGACCCGCCCTGGCCCGAGGTGTCCCA 266  
 QY 155 AlaCysAlaGlyArgPheGlnSerProValAspIleArgProGlnLeuAlaAlaPheCys 174  
 DB 267 GCCTGGCGGGCCCTTCCAGTCCCGGTAGATATCCGCCCCAGCTCGCCGCTTCTGC 326  
 QY 175 ProAlaLeuArgProLeuGluLeuGlyPheGlnLeuProProLeuProGluLeuArg 194  
 DB 327 CCCGCCCTGGACCCCTGGAACTCTCTGGGCTTCGAGCTCCCGCTGCTCCAGAACTGGCG 386

QY 195 LeuArgAsnAsnGlyHisSerValGlnLeuThrLeuProProGlyLeuGluMetAlaLeu 214  
 DB 387 CTGCGCAACAATGGCCACAGTGTGAACCTGACCTGCCCTCCGGGCTAGAGATGGCTCTG 446  
 QY 215 GlyProGlyArgGluTyrArgAlaLeuGlnLeuHisLeuHisTrpGlyAlaAlaGlyArg 234  
 DB 447 GGTCCCGCGGGGAGTACCGGGCTCTGCAGCTGCATCTGCACCTGGGGTCTGTAGGTCTGT 506  
 QY 235 ProGlySerGluHisThrValGluGlyHisArgPheProAlaGluIleHisValValHis 254  
 DB 507 CCGGGCTCGAGACACACTGTGGAAGGCCACCGCTTTCCCTGCCGAGATCCACCGTGTTCAC 566  
 QY 255 LeuSerThrAlaPheAlaArgValAspGluAlaLeuGlyArg-ProGlyGlyLeuAlaVa 274  
 DB 567 ATCAGACCCGCTATGCCAGAGTTGGAGGGCTTTGGGGCGCCCCCGAGGCCCTGGCCGT 626  
 QY 274 lLeuAlaAlaPheLeuGluGluGlyProGluGluAsnSerAlaTyrGluGlnLeuLeuSe 294  
 DB 627 GTTGGCGGCTTTCTGGAGGAGGCCCGGAGAAACACAGTGCCTATGAGCAGTTGCTGTC 686  
 QY 294 rArgLeuGluGluIleAlaGluGlySerGluThrGlnValProGlyLeuAspIleSe 314  
 DB 687 TCACCTTGGAGAAATCGCTGAGAAAGGCTCAGAGACTCAGTCCCAAGGACTGGACATATC 746  
 QY 314 rAlaLeuLeuProSerAspPheSerArgTyrPheGlnTyrGluGlySerLeuThrPr 334  
 DB 747 TGCACCTCTCCCTCTGACCTCAGCCGCTACTTCGATATGAGGGGTCTCTGACTACAG 806  
 QY 334 oProCysAlaGlnGlyValIleTrp-ThrValPheAsnGlnThrValMetLeuSerAlaL 354  
 DB 807 GCCTGTGTCAGGGGTGTCATCTGGGACGTGTGTTTTACCAGACAGTGTGCTGAGTGCTA 866  
 QY 354 yGlnLeuHisThrLeuSerAspThrLeu-TrpGlyProGlyAspSer-ArgLeuGlnLe 373  
 DB 867 ACCACCTCCACCCCTCTCTGAAGCCCTGGGGGGACCTGGGGACTCTCCGGCTACGGCT 926  
 QY 373 uAsnPhe---ArgAlaThrGlnProLeuAsnGly 383  
 DB 927 GAACCTTCGAGAGGGAGCCACCCTTGGGAATGGG 960  
 RESULT 13  
 LOCUS AL555184  
 DEFINITION AL555184 Homo sapiens HELA CELLS COT 25-NORMALIZED Homo sapiens  
 CDNA clone CS0DK007YK10 5-PRIME, mRNA sequence.  
 ACCESSION AL555184  
 VERSION AL555184.3 GI:45859924  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
 Homnidae; Homo.  
 REFERENCE 1 (bases 1 to 927)  
 AUTHORS Li, W.B., Gruber, C., Jesse, J., and Polayes, D.  
 TITLE Full-length cDNA libraries and normalization  
 JOURNAL Unpublished (2001)  
 COMMENT On Feb 15, 2001 this sequence version replaced gi:31276993.  
 Contact: Genoscope  
 Genoscope - Centre National de Sequencage  
 2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE  
 Email: segref@genoscope.cns.fr, Web : www.genoscope.cns.fr  
 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime  
 end enriched, double-strand cDNA was digested with Not I and cloned  
 into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library  
 was normalized. Library was constructed by Life Technologies, a  
 division of Invitrogen. This sequence belongs to sequence cluster  
 5300.f  
 For more information about this cluster, see  
 http://www.genoscope.cns.fr/cdna?8=CS0DK007BF050QP1&c=5300.f.  
 FEATURES  
 source Location/Qualifiers  
 1. .927  
 /organism="Homo sapiens"





```
Db      662 CTGAGTCTAAGCAGCTCCACACCCTCTCTGACTCCCTGTGGGACCTGATGACTCTCGG 721
Qy      371 LeuGlnLeuAsnPheArgAlaThrGlnProLeuAsnGlyArgValIleGluAlaSerPhe 390
Db      722 CTGCAGCTGAACCTCCGAGCTACGCAGCCTTTGAATGGGCGAATAATTGAGGCTCCTTC 781
Qy      391 ProAlaGlyValAspSerSerProArgAlaAlaGluProValGlnLeuAsnSerCysLeu 410
Db      782 CCCGCTGGCGTGGATGGCAGCCCTAGGACTGTGTGACACGATCCACCTGAAATTCCTGTCTC 841
Qy      411 AlaAlaGlyAspIleLeuAlaLeuValPheGlyLeuLeuPheAla 425
Db      842 GCTGCTGGCGACATCCTGNCCTGGTTTGGGCTCCTCTTTGCT 886
```

Search completed: February 17, 2006, 18:21:44  
Job time : 5591 secs

**THIS PAGE BLANK (uapftr)**

GenCore version 5.1.7  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: February 17, 2006, 15:10:31 ; Search time 316 Seconds  
(without alignments)

2581.965 Million cell updates/sec

Title: US-09-967-237B-2

Perfect score: 2424

Sequence: 1 MAPLCPSPWLLIPAPAPG.....RRGTGGVSRPAEVAETGA 459

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 1303057 seqs, 888780828 residues

Total number of hits satisfying chosen parameters: 2606114

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%  
Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlip  
-Q=/abs/ABSSWEB spool/US09967237/runat\_15022006\_182246\_28055/app\_query.fasta\_1  
-DB=Issued Patents NA -OPMT=fastap -SUFFIX=rni -MINMATCH=0.1 -LOOPL=0  
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi  
-LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15  
-MODE=LOCAL -OUTFMT=ptc -NORM=seq -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000  
-HOST=abs06p -USER=US09967237 @CGN 1 1 290 @runat\_15022006\_182246\_28055  
-NCPU=6 -ICPU=3 -NO MMAP -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOPOP=6  
-FGAPEXT=7 -YGAPOP=10 -XGAEXT=0.5 -DELOP=6 -DELEXT=7

Database : Issued Patents NA.\*

1: /cgn2\_6/ptodata/1/ina/1 COMB.seq.\*  
2: /cgn2\_6/ptodata/1/ina/5 COMB.seq.\*  
3: /cgn2\_6/ptodata/1/ina/6A COMB.seq.\*  
4: /cgn2\_6/ptodata/1/ina/6B COMB.seq.\*  
5: /cgn2\_6/ptodata/1/ina/H COMB.seq.\*  
6: /cgn2\_6/ptodata/1/ina/PCTUS COMB.seq.\*  
7: /cgn2\_6/ptodata/1/ina/PP COMB.seq.\*  
8: /cgn2\_6/ptodata/1/ina/RE COMB.seq.\*  
9: /cgn2\_6/ptodata/1/ina/backfiles1.seq.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2424	100.0	1522	2	US-08-481-658B-1
2	2424	100.0	1522	2	US-08-477-504A-1
3	2424	100.0	1522	2	US-08-486-756A-1
4	2424	100.0	1522	2	US-08-485-862B-1
5	2424	100.0	1522	3	US-08-787-739-1
6	2424	100.0	1522	3	US-08-487-077A-1
7	2424	100.0	1522	3	US-08-485-863A-1
8	2424	100.0	1522	3	US-08-485-049D-1
9	2424	100.0	1522	3	US-09-178-115-1

10	2424	100.0	1522	3	US-09-177-776-1	Sequence 1, Appli
11	2424	100.0	1522	3	US-09-772-719B-1	Sequence 1, Appli
12	2424	100.0	1522	3	US-08-260-190-5	Sequence 5, Appli
13	2424	100.0	1522	3	US-09-949-016-171	Sequence 171, App
14	2424	99.8	1522	3	US-09-949-016-4332	Sequence 4332, Ap
15	2233	92.1	5052	3	US-08-260-190-23	Sequence 23, Appli
16	2227	91.9	1399	3	US-08-335-469-1	Sequence 1, Appli
17	2227	91.9	1399	3	US-08-260-190-1	Sequence 1, Appli
18	2036	84.0	1397	6	US-07-964-589-1	Sequence 1, Appli
19	2036	84.0	1397	6	PCT-US93-02024-1	Sequence 11913, A
20	1253.5	51.7	11237	3	US-09-949-016-11913	Sequence 16074, A
21	1253.5	51.7	11237	3	US-09-949-016-16074	Sequence 5, Appli
22	1176.5	48.5	10898	2	US-08-481-658B-5	Sequence 5, Appli
23	1176.5	48.5	10898	2	US-08-477-504A-5	Sequence 5, Appli
24	1176.5	48.5	10898	2	US-08-486-756A-5	Sequence 5, Appli
25	1176.5	48.5	10898	2	US-08-485-862B-5	Sequence 5, Appli
26	1176.5	48.5	10898	3	US-08-787-739-5	Sequence 5, Appli
27	1176.5	48.5	10898	3	US-08-487-077A-5	Sequence 5, Appli
28	1176.5	48.5	10898	3	US-08-485-863A-5	Sequence 5, Appli
29	1176.5	48.5	10898	3	US-08-485-049D-5	Sequence 5, Appli
30	1176.5	48.5	10898	3	US-09-178-115-5	Sequence 5, Appli
31	1176.5	48.5	10898	3	US-09-177-776-5	Sequence 5, Appli
32	1176.5	48.5	10898	3	US-09-772-719B-5	Sequence 5, Appli
33	714	29.5	415	2	US-08-481-658B-28	Sequence 28, Appli
34	714	29.5	415	2	US-08-477-504A-28	Sequence 28, Appli
35	714	29.5	415	2	US-08-486-756A-28	Sequence 28, Appli
36	714	29.5	415	2	US-08-485-862B-28	Sequence 28, Appli
37	714	29.5	415	3	US-08-487-077A-28	Sequence 28, Appli
38	714	29.5	415	3	US-08-485-863A-28	Sequence 28, Appli
39	714	29.5	415	3	US-08-485-049D-28	Sequence 28, Appli
40	714	29.5	445	3	US-08-787-739-28	Sequence 28, Appli
41	714	29.5	445	3	US-09-178-115-28	Sequence 28, Appli
42	714	29.5	445	3	US-09-177-776-28	Sequence 28, Appli
43	714	29.5	445	3	US-09-772-719B-28	Sequence 28, Appli
44	629	25.9	1401	2	US-08-481-658B-49	Sequence 49, Appli
45	629	25.9	1401	2	US-08-477-504A-49	Sequence 49, Appli

#### ALIGNMENTS

#### RESULT 1

US-08-481-658B-1

; Sequence 1, Application US/08481658B

; Patent No. 5955075

; GENERAL INFORMATION:

; APPLICANT: Zavada, Jan

; APPLICANT: Pastorekova, Silvia

; APPLICANT: Pastorek, Jaromir

; TITLE OF INVENTION: MN Gene and Protein

; NUMBER OF SEQUENCES: 86

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Leona L. Lauder

; STREET: 6 Mariposa Court

; CITY: Tiburon

; STATE: California

; COUNTRY: USA

; ZIP: 94920

; COMPUTER READABLE FORM: disk

; MEDIUM TYPE: Floppy

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patent In Release #1.0, Version #1.30 (BPO)

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/481,658B

; FILING DATE: 07-JUN-1995

; CLASSIFICATION: 424

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 08/260,190

; FILING DATE: 15-JUN-1994

; ATTORNEY/AGENT INFORMATION:

; NAME: Lauder, Leona L.

; REGISTRATION NUMBER: 30,863

; REFERENCE/DOCKET NUMBER: D-0021.3E

## TELECOMMUNICATION INFORMATION:

TELEPHONE: 415-435-2034  
TELEFAX: 415-435-0727  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1522 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
US-08-481-658B-1

## Alignment Scores:

Pred. No.: 7, 5e-197 Length: 1522  
Score: 2424.00 Matches: 459  
Percent Similarity: 100.0% Conservativeness: 0  
Best Local Similarity: 100.0% Mismatches: 0  
Query Match: 100.0% Indels: 0  
DB: 2 Gaps: 0

US-09-967-237B-2 (1-459) x US-08-481-658B-1 (1-1522)

QY 1 MetAlaProLeuCysProSerProTrpLeuProLeuLeuLeuProAlaProGly 20  
DB 13 ATGGCTCCCTGTGCGCCAGCCCTGGCTCCCTCTGTGATCCCGGCCCTGTCTCCAGGC 72  
QY 21 LeuThrValGlnLeuLeuLeuSerLeuLeuLeuMetProValHisProGlnArgLeu 40  
DB 73 CTCACTGTGCACTGCTGTGTCACCTGCTGCTTCTGATGCTGTCCATCCCCAGAGTTG 132  
QY 41 ProArgMetGlnGluAspSerProLeuGlyGlyGlySerSerGlyGluAspProLeu 60  
DB 133 CCCCGATGACAGAGGATCCCCCTTTGGGAGGAGGCTCTTCTGGGAAGATGATGACCCACTG 192  
QY 61 GlyGluGluAspLeuProSerGluGluAspSerProArgGluGluAspProProGlyGlu 80  
DB 193 GGCAGAGGAGATCTGCCCAATGAGAGGATTCACCCAGAGAGGAGATCCACCCCGAGAG 252  
QY 81 GluAspLeuProGlyGluGluAspLeuProGlyGluGluAspLeuProGluValLysPro 100  
DB 253 GAGATCTACCTGGAGGAGGATCTACCTGGAGAGGAGATCTCTGAGGTCTCTGGA 312  
QY 101 LysSerGluGluGlySerLeuLysLeuGluAspLeuProThrValGluAlaProGly 120  
DB 313 AATCAGAGAGAGGCTCCCTGAAAGTTAGAGGATCTACTACTGTTGAGGCTCTCTGGA 372  
QY 121 AspProGlnGluProGlnAsnAsnAlaHisArgAspLysGluGlyAspAspGlnSerHis 140  
DB 373 GATCCTCAAGAACCCAGATAATGCCACAGGGACAAAGAGGGATGACCAAGTCTAT 432  
QY 141 TrpArgTrpGlyGlyAspProProTrpProArgValSerProAlaCysAlaGlyArgPhe 160  
DB 433 TGGCGCTATGAGCGAGACCCGCTGCGCCCGGGTGTCCACAGCTTGGCGGGCGCGCTTC 492  
QY 161 GlnSerProValAspLeuArgProGlnLeuAlaAlaPheCysProAlaLeuArgProLeu 180  
DB 493 CAGTCCCGGTGTATATCCGCCCCAGCTGCGCGCTTCTGCGCGGCTGCGCGCCCTG 552  
QY 181 GluLeuLeuGlyPheGlnLeuProProLeuProGluLeuArgLeuArgAsnGlyHis 200  
DB 553 GAATCTCTGGGCTTCCAGCTCCCGCGCTCCCAAGAACTGGCGCTGGCGCAATGGCCAC 612  
QY 201 SerValGlnLeuThrLeuProProGlyLeuGluMetAlaLeuGlyProGlyArgGluTrp 220  
DB 613 AGTGTCAACTGACCTGCTCTGGGTAGAGATGGCTCTGGGTCCCGGGCGGGAGTAC 672  
QY 221 ArgAlaLeuGlnLeuHisLeuHisTrpGlyAlaAlaGlyArgProGlySerGluHisThr 240  
DB 673 CCGGCTCTGAGCTGATCTGCACTGGGGGCTGCAGGTGCTCCGGGCTCGGAGCACT 732  
QY 241 ValGluGlyHisArgPheProAlaGluLeuHisValValHisLeuSerThrAlaPheAla 260

DB 733 GTGGAAGGCCACCGTTTCCCTCCAGATCCACGGTTCACCTCAGCACCCCTTTGGC 792  
QY 261 ArgValAspGluAlaLeuGlyArgProGlyGlyLeuAlaValLeuAlaPheLeuGlu 280  
DB 793 AGAGTTGACGAGGCTTGGGGCGCCGGAGGCTGGCGGTGTGGCCGCTTTCTGGAG 852  
QY 281 GluGlyProGluGluAsnSerAlaTrpGluGlnLeuLeuSerArgLeuGluLeuAla 300  
DB 853 GAGGGCCCGAAGAAAACAGTGCCTATGAGCAGTTGCTGTCTCGCTTGAAGAAATCGCT 912  
QY 301 GluGluGlySerGluThrGlnValProGlyLeuAspIleSerAlaLeuLeuProSerAsp 320  
DB 913 GAGGAAGGCTCAGAGACTCAGGTCCAGACTGGACATATCTGCATCTCGCCCTCTGAC 972  
QY 321 PheSerArgTrpPheGlnTrpGlySerLeuThrProProCysAlaGlnGlyVal 340  
DB 973 TTCAGCGCTACTTCCAAATATGAGGGTCTCTGACTACACCGCCCTGTGCCAGGGTGC 1032  
QY 341 IleTrpThrValPheAsnGlnThrValMetLeuSerAlaLysGlnLeuHisThrLeuSer 360  
DB 1033 ATCTGGACTGTCTTTAAACCAAGACAGTGTGCTGAGTGTCTTAAGCAGCTCCACACCTCTCT 1092  
QY 361 AspThrLeuTrpGlyProGlyAspSerArgLeuGlnLeuAsnPheArgAlaThrGlnPro 380  
DB 1093 GACACCTGTGGGACCTGGTACTCTCGGCTACAGCTGACTTCGAGCGGAGCGCT 1152  
QY 381 LeuAsnGlyArgValIleGluAlaSerPheProAlaGlyValAspSerSerProArgAla 400  
DB 1153 TTGAATGGSCGAGTGATTGAGGCTCTTCTCCCTGCTGGAGTGGACAGCAGTCTCTCGGCT 1212  
QY 401 AlaGluProValGlnLeuAsnSerCysLeuAlaAlaGlyAspIleLeuAlaLeuValPhe 420  
DB 1213 GCTGAGCCAGTCCAGCTGAATTCCTGCTGGCTGCTGGTGTGATCTCTAGCCCTGGTTTT 1272  
QY 421 GlyLeuLeuPheAlaValThrSerValAlaPheLeuValGlnMetArgArgGlnHisArg 440  
DB 1273 GGCCTCTTTTGTGTCTACAGCGTCCGCTTCTTGTGCGATGTGCGATGAGAGGAGCAGACA 1332  
QY 441 ArgGlyThrLysGlyValSerTrpArgProAlaGluValAlaGluThrGlyAla 459  
DB 1333 AGGGGAACCAAGGGGTGTGAGCTACCGCCAGAGAGGTAGCCGAGACTGGAGCC 1389

## RESULT 2

US-08-477-504A-1  
Sequence 1, Application US/08477504A  
Patent No. 5972353  
GENERAL INFORMATION:  
APPLICANT: Zavada, Jan  
APPLICANT: Pastorekova, Silvia  
APPLICANT: Pastorek, Jaromir  
TITLE OF INVENTION: MN Gene and Protein  
NUMBER OF SEQUENCES: 86  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Leona L. Lauder  
STREET: 6 Mariposa Court  
CITY: Tiburon  
STATE: California  
COUNTRY: USA  
ZIP: 94920  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.30 (EPO)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/477,504A  
FILING DATE: 07-JUN-1995  
CLASSIFICATION: 424  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/260,190  
FILING DATE: 15-JUN-1994  
ATTORNEY/AGENT INFORMATION:

NAME: Lauder, Leona L.  
REGISTRATION NUMBER: 30,863  
REFERENCES/DOCKET NUMBER: D-0021.3D  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 415-435-2034  
TELEFAX: 415-435-0727  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1522 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
US-08-477-504A-1

Alignment Scores:  
Pred. No.: 7,5e-197 Length: 1522  
Score: 2424.00 Matches: 459  
Percent Similarity: 100.0% Conservative: 0  
Best Local Similarity: 100.0% Mismatches: 0  
Query Match: 100.0% Indels: 0  
DB: 2 Gaps: 0

US-09-967-237B-2 (1-459) x US-08-477-504A-1 (1-1522)

QY	1	MetAlaProLeuCysProSerProTrrLeuProLeuLeuIleProAlaProAlaProGly	20
DB	13	ATGGCTCCCTCTGCCCCAGCCCCCTGGCTCCCTCTGTTGATCCCCGGGCCCCCTGCCAGGC	72
QY	21	LeuThrValGlnLeuLeuSerLeuLeuLeuMetProValHisProGlnArgLeu	40
DB	73	CTCATCTGTCACTGCTGTCTGCTCTGCTCTGATGCCCTGCCATCCCCAGAGTTG	132
QY	41	ProArgMetGlnGluAspSerProLeuGlyGlySerSerGlyGluAspAspProLeu	60
DB	133	CCCCGGATCAGAGAGATTCCCTCTGGGAGAGGCTCTTCTGGGGAAGATGACCCACTG	192
QY	61	GlyGluGluAspLeuProSerGlnGluAspSerProArgGluGluAspProProGlyGlu	80
DB	193	GGCGAGAGAGATCTGCCAGTGAAGAGATTCCACCAGAGAGAGAGATCCACCCTGGAGAG	252
QY	81	GluAspLeuProGlyGluGluAspLeuProGlyGluGluAspLeuProGluValLeuPro	100
DB	253	GAGGATCTACTTGGAGAGAGGATCTACTCTGGAGAGAGGATCTACTGAAAGTTAAGCCT	312
QY	101	LysSerGlnGluGlySerLeuLysLeuGluAspLeuProThrValGluAlaProGly	120
DB	313	AAATCAGAGAGAGGGCTCCCTGAAAGTTAGAGGATCTACTCTGTTGAGGCTCCCTGGA	372
QY	121	AspProGlnGluProGlnAsnAsnAlaHisArgAspLysGluGlyAspAspGlnSerHis	140
DB	373	GATCTCTAGAACCCACAGATTAATGCCACAGGGAACAAGAGGGGATGACCAGAGTCAT	432
QY	141	TrpArgTyrGlyCysAspProProTrpProArgValSerProAlaCysAlaGlyArgPhe	160
DB	433	TGGCGCTATGAGGGACACCGCCCTGGCCCCGGGTGTCCTCCAGCGCTGGCGGGCGGCTTC	492
QY	161	GlnSerProValAspIleArgProGlnLeuAlaPheCysProAlaLeuArgProLeu	180
DB	493	CAGTCCCCGGTGATATCCGCCCCCAGCTCGCGGCTCTCTGCCCCGGGCTCTGCCCCCTG	552
QY	181	GluLeuLeuGlyPheGlnLeuProProLeuProGluLeuArgLeuArgAsnAsnGlyHis	200
DB	553	GAACTCTGGGCTTCCAGCTCCCGCGCTCCAGAACTCCGCTGGCGCAACAATGGCCAC	612
QY	201	SerValGlnLeuThrLeuProProGlyLeuGluMetAlaLeuGlyProGlyArgGluTyr	220
DB	613	AGTGTCAACTGACCTGCTCTGGGCTAGAGATGGCTCTGGGTGCCGGCGGGAGTAC	672
QY	221	ArgAlaLeuGlnLeuHisLeuHisTrpGlyAlaAlaGlyArgProGlySerGluHisThr	240









QY 441 ArgGlyThrIysGlyGlyValSerTyrArgProAlaGluValAlaGluThrGlyAla 459  
Db 1333 AGGGGAACCAAGGGGTGTGAGCTACCGCCAGCAGAGGTAGCGAGCTGGAGCC 1389

## RESULT 7

US-08-485-863A-1  
; Sequence 1, Application US/08485863A  
; Patent No. 6093548  
; GENERAL INFORMATION:  
; APPLICANT: Zavada, Jan  
; APPLICANT: Pastorekova, Silvia  
; APPLICANT: Pastorek, Jaromir  
; TITLE OF INVENTION: MN Gene and Protein  
; NUMBER OF SEQUENCES: 86  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Leona L. Lauder  
; STREET: 6 Mariposa Court  
; CITY: Tiburon  
; STATE: California  
; COUNTRY: USA  
; ZIP: 94920  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.30 (BPO)  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/485,863A  
; FILING DATE: 07-JUN-1995  
; CLASSIFICATION: 514  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/260,190  
; FILING DATE: 15-JUN-1994  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Lauder, Leona L.  
; REGISTRATION NUMBER: 30,863  
; REFERENCE/DOCKET NUMBER: D-0021.3G  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 415-435-2034  
; TELEFAX: 415-435-0727  
; INFORMATION FOR SEQ ID NO: 1:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1522 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: cDNA  
; HYPOTHEetical: NO  
; ANTI-SENSE: NO  
; US-08-485-863A-1

Alignment Scores:  
Pred. No.: 7,5e-197 Length: 1522  
Score: 2424.00 Matches: 459  
Percent Similarity: 100.0% Conservative: 0  
Best Local Similarity: 100.0% Mismatches: 0  
Query Match: 100.0% Indels: 0  
DB: 3 Gaps: 0

US-09-967-237B-2 (1-459) x US-08-485-863A-1 (1-1522)

QY 1 MetAlaProLeuCysProSerProTyrProLeuProLeuLeuLeuProAlaProAlaProGly 20  
Db 13 ATGGCTCCCTGTGCCCCAGCCCTGGCTCCCTCTGTGTGATCCCGGCCCTGTCTCCAGGC 72  
QY 21 LeuThrValGlnLeuLeuLeuSerLeuLeuLeuMetProValHisProGlnArgLeu 40  
Db 73 CTCACTGTGCACTGCTGCTGCTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 132  
QY 41 ProArgMetGlnGluAspSerProLeuGlyGlySerSerGlyGluAspProLeu 60  
Db 133 CCCCGATGTCAGGAGGATTCCTCCCTTGGAGGAGGCTCTTCTTGGGAAGATGACCCACTG 192

QY 61 GlyGluGluAspLeuProSerGluGluAspSerProArgGluGluAspProProGlyGlu 80  
Db 193 GCGAGGAGGATCTGCCAGTAGAGGATTCACCCAGAGAGGAGATCCACCCGAGAG 252  
QY 81 GluAspLeuProGlyGluGluAspLeuProGlyGluGluAspLeuProGluValLysPro 100  
Db 253 GAGGATCTACCTGGAGAGGAGATCTACCTGGAGAGGAGATCTACTGAGGTAGACCT 312  
QY 101 LysSerGluGluGlySerLeuLysLeuGluAspLeuProThrValGluAlaProGly 120  
Db 313 AAATCAGAAAGAGAGGCTCCCTGAAAGTTAGAGGATCTACTTCTGAGGCTCTCGGA 372  
QY 121 AspProGlnGluProGlnAsnAlaHisArgAspLysGluGlyAspGlnSerHis 140  
Db 373 GATCTCTCAAGAACCCAGGAATAATGCCACACAGGACAAAGAGGAGATGACCAAGTGCAT 432  
QY 141 TpaArgTyrGlyGlyAspProProTyrProArgValSerProAlaCysAlaGlyArgPhe 160  
Db 433 TGGCGCTATGGAGGCGACCCGCTTGGCCCGGGTGTCCCGAGCTTGGCGGGGCGCTTC 492  
QY 161 GlnSerProValAspIleArgProGlnLeuAlaAlaPheCysProAlaLeuArgProLeu 180  
Db 493 CAGTCCCCGGTGGATATCCGCCCCAGCTCGCCGCTTCTGCCCGGCTCGCGCCCTG 552  
QY 181 GluLeuLeuGlyPheGlnLeuProProLeuProGluLeuLeuArgLeuArgAsnGlyHis 200  
Db 553 GAACTCTGGGCTTCAGCTCCCGCTCCCAAGAACTGCGCTGCGCAACAATGGCCAC 612  
QY 201 SerValGlnLeuThrLeuProProGlyLeuGluMetAlaLeuGlyProGlyArgGluTyr 220  
Db 613 AGTGTGCACTGACCTGCTCTGGGCTAGAGTGGCTCTGGTCCCGGGGAGTAC 672  
QY 221 ArgAlaLeuGlnLeuHisLeuHisTrpGlyAlaAlaGlyArgProGlySerGluHisThr 240  
Db 673 CGGGCTCTGCAGCTGCATCTGCACCTGGGGGCTGCAGGTCGTCCGGGCTCGGAGCACACT 732  
QY 241 ValGluGlyHisArgPheProAlaGluLleHisValHisLeuSerThrAlaPheAla 260  
Db 733 GTGGAAGGCCACCGTTTCCCTGCCAGATCCACGTGGTTCACCTCAGCACCGCTTTGCC 792  
QY 261 ArgValAspGluAlaLeuGlyArgProGlyGlyLeuAlaValLeuAlaAlaPheLeuGlu 280  
Db 793 AGAGTTGACGAGGCTTGGGGGCGCCGGGAGGCGCTGGCCGTGTGGCCGCTTTCTGGAG 852  
QY 281 GluGlyProGluGluAsnSerAlaTyrGluGlnLeuLeuSerArgLeuGluLleAla 300  
Db 853 GAGGGCCCGAAGAAAACAGTGCCTATGAGCAGTTGCTGCTCGCTTGGAAAGAAATCGCT 912  
QY 301 GluGluGlySerGlnThrGlnValProGlyLeuAspIleSerAlaLeuLeuProSerAsp 320  
Db 913 GAGGAAGGCTCAGAGACTCAGGTCCCGAGGACTGGACATATCTGCATCTCTGCCCTCTGAC 972  
QY 321 PheSerArgTyrPheGlnTyrGluGlySerLeuThrThrProCysAlaGlnGlyVal 340  
Db 973 TTCAGCGGCTACTTCCCAATATGAGGGTCTCTGACTACACCGCCCTGTGCCAGGGGTGC 1032  
QY 341 IleTyrThrValPheAsnGlnThrValMetLeuSerAlaLysGlnLeuHisThrLeuSer 360  
Db 1033 ATCTGAGCTGTGTTAAACCAGACAGTGTGCTGAGTGCTAAGCAGGTCCACACCTCTCT 1092  
QY 361 AspThrLeuTyrProGlyAspSerArgLeuGlnLeuAsnPheArgAlaThrGlnPro 380  
Db 1093 GACACCTGTGGGAGACTGTGTGACTCTCGGCTACAGCTGACTTCGAGGAGCGAGCCT 1152  
QY 381 LeuAsnGlyArgValIleGluAlaSerPheProAlaGlyValAspSerSerProArgAla 400  
Db 1153 TTGAATGGGCGAGTGTGAGGCTCTCTTCCCTGCTGGAGTGGAGCAGCAGTCTCTCGGCT 1212  
QY 401 AlaGluProValGlnLeuAsnSerCysLeuAlaAlaGlyAspIleLeuAlaLeuValPhe 420  
Db 1213 GCTGAGCCAGTCCAGCTGAATTCCTGCTGGCTGGTGGTGCATCTCCTAGCCTGGTGTCT 1272  
QY 421 GlyLeuLeuPheAlaValThrSerValAlaPheLeuValGlnMetArgArgGlnHisArg 440

```

Db      1273  GGCTCTCTTTGCTGTCCAGCGTCCGCTCTCTTGTGCAGATGAGAGCGACACAGA 1332
Qy      441  ArgGlyThrIysGlyGlyValSerTyrArgProAlaGluValAlaGluThrGlyAla 459
Db      1333  AGGGGAACCAAGGGGTGTGAGCTACCGCCAGCAGAGGTAGCCGAGACTGGAGCC 1389

RESULT 8
US-08-485-049D-1
; Sequence 1, Application US/08485049D
; Patent No. 6204370
; GENERAL INFORMATION:
; APPLICANT: Zavada, Jan
; APPLICANT: Pastorekova, Silvia
; APPLICANT: Pastorek, Jaromir
; TITLE OF INVENTION: MN Gene and Protein
; NUMBER OF SEQUENCES: 86
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Leona L. Lauder
; STREET: 369 Pine Street
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30 (BPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/485,049D
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/260,190
; FILING DATE: 15-JUN-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Lauder, Leona L.
; REGISTRATION NUMBER: 30,863
; REFERENCE/DOCKET NUMBER: D-0021.3E
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-981-2034
; TELEFAX: 415-981-0332
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1522 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; US-08-485-049D-1

Alignment Scores:
Pred. No.:      7 5e-157      Length:      1522
Score:          2424.00      Matches:      459
Percent Similarity: 100.0%      Conservative: 0
Best Local Similarity: 100.0%      Mismatches: 0
Query Match:      100.0%      Indels:      0
DB:              3          Gaps:      0

US-09-967-237B-2 (1-459) x US-08-485-049D-1 (1-1522)
Qy      1  MetAlaProLeuCysProSerProTrrProLeuProLeuLeuLeuProAlaProGly 20
Db      13  ATGGCTCCCTGTGCGCCAGCCCTGGCTCCCTGTGTGATCCCGGCCCTGTCTCCAGGC 72
Qy      21  LeuThrValGlnLeuLeuLeuSerLeuLeuLeuMetProValHisProGlnArgLeu 40
Db      73  CTCAGTGTCACTGTGCTGTCTCACTGCTCTTGTGATGCTGTCTCACTCCAGAGGTTG 132
Qy      41  ProArgMetGlnGluAspSerProLeuGlyGlySerSerGlyGluAspProLeu 60

```

```

133  CCCCAGATCAGAGGATTTCCCTTTGGAGGAGGCTCTTCTGGGAGAGATGACCCACTG 192
Qy      61  GlyGluGluAspLeuProSerGluGluAspSerProArgGluGluAspProProGlyGlu 80
Db      193  GCGAGAGAGGATCTGCCAGTGAAGAGGATTTCAACCAGAGAGGAGATCCACCCGAGAG 252
Qy      81  GluAspLeuProGlyGluGluAspLeuProGlyGluGluAspLeuProGluValLysPro 100
Db      253  GAGGATCTACCTGGAGAGGAGGATCTACCTGGAGAGGAGGATCTACCTGAAGTTAAGCCT 312
Qy     101  LysSerGluGluGluGlySerLeuLysLeuGluAspLeuProThrValGluAlaProGly 120
Db     313  AAATCAGAAAGAGAGGCTCCCTGAAGTTAGAGGATCTACCTACTGTTGAGGCTCTCGGA 372
Qy     121  AspProGlnGluProGlnAenAenAlaHisArgAspLysGluGlyAspAspGlnSerHis 140
Db     373  GATCCTCAAGAACCCAGAAATATATGCCACAGGAGCAAGAGAGGATGACAGAGTCAT 432
Qy     141  TrpArgTyrGlyGlyAspProProTrrProArgValSerProAlaCysAlaGlyArgPhe 160
Db     433  TGGCGCTATGGAGGCGACCCGCCCTGGCCCGGGTGTCCCCAGCCTGGCGCGCGCTTC 492
Qy     161  GlnSerProValAspLysArgProGlnLeuAlaAlaPheCysProAlaLeuArgProLeu 180
Db     493  CAGTCCCGGTGGATATCGCGCCCGCCAGCTCGCGCGCTTCTGCGCGCGCTCGCGCCCTG 552
Qy     181  GluLeuLeuGlyPheGlnLeuProProLeuProGluLeuArgLeuArgLeuAenGlyHis 200
Db     553  GAACTCTGGGCTTCAGCTCCCGCGCTCCCAAGAACTCGCGCTCGCGCAACATGGCCAC 612
Qy     201  SerValGlnLeuThrLeuProProGlyLeuGluMetAlaLeuGlyProGlyArgGluTyr 220
Db     613  AGTGTGCAACTGACCTGCTCTGCGCTAGAGATGGCTCTGGGTCCCGGCGGAGTAC 672
Qy     221  ArgAlaLeuGlnLeuHisLeuHisTrrGlyAlaAlaGlyArgProGlySerGluHisThr 240
Db     673  CGGGCTCTGCAGCTGCATCTGCACCTGGGGGCTGCAGGTCGTCCGGCTCGGACACACT 732
Qy     241  ValGluGlyHisArgPheProAlaGluLeuHisValValHisLeuSerThrAlaPheAla 260
Db     733  GTGGAAGGCCACCGTTTCCCTGCGAGATCCACGTGGTTCCACTCAGCACCGCTTTGGCC 792
Qy     261  ArgValAspGluAlaLeuGlyArgProGlyGlyLeuAlaValLeuAlaAlaPheLeuGlu 280
Db     793  AGAGTTGACAGGCTTTGGGGCGCCCGGAGGCTCGCGCGTGTGGCCGCTTCTCGAG 852
Qy     281  GluGlyProGluGluAenSerAlaTyrGluGlnLeuLeuSerArgLeuGluLileAla 300
Db     853  GAGGGCCCGGAAGAAAACAGTGCTATGAGCAGTTGCTGTCTCGCTTGGGAAGAAATCGCT 912
Qy     301  GluGluGlySerGlnThrGlnValProGlyLeuAspLysSerAlaLeuLeuProSerAsp 320
Db     913  GAGGAAGGCTCAGAGACTCAGGTCCAGGATCCGACATATCTGCATCTCTGCCCTCTGAC 972
Qy     321  PheSerArgTyrPheGlnTyrGluGlySerLeuThrThrProProCysAlaGlnGlyVal 340
Db     973  TTCAGCGGCTACTTCCAAATATGAGGGGTCTCTGACTACACCGCCCTGTGCCAGGGTGT 1032
Qy     341  IleTrrThrValPheAenGlnThrValMetLeuSerAlaLysGlnLeuHisThrLeuSer 360
Db    1033  ATCTGAGCTGTGTTAAACAGACAGTGATGCTGAGTGCTAAGCAGCTCCACACCTCTCT 1092
Qy     361  AspThrLeuTrrGlyProGlyAspSerArgLeuGlnLeuAenPheArgAlaThrGlnPro 380
Db    1093  GACACCTGTGGGAGACTCTGTGACTCTGGCTACAGCTGAACCTCCGAGGAGCGACGCT 1152
Qy     381  LeuAenGlyArgValIleGluAlaSerPheProAlaGlyValAspSerSerProArgAla 400
Db    1153  TTGAATGGGCGAGTGATTTAGGCTCTCTTCCCTGCTGGAGTGGACAGCAGCTCTCGGGCT 1212
Qy     401  AlaGluProValGlnLeuAenSerCysLeuAlaAlaGlyAspIleLeuAlaLeuValPhe 420

```

Db 1213 GCTGAGCCAGTCCAGCTGAATTCCTGCGCTGGTGCATCTCCTAGCCCTGGTGTTC 1272  
Qy 421 GlyLeuLeuPheAlaValThrSerValAlaPheLeuValGlnMetArgArgGlnHisArg 440  
Db 1273 GGCCTCCTTTTGTGTGCACAGCGTCCGCTTCCTTGTGCAGATGAGAGGCAGACAGA 1332  
Qy 441 ArgGlyThrIysGlyGlyValSerTyrArgProAlaGluValAlaGluThrGlyAla 459  
Db 1333 AGGGAAACCAAGGGGTGTGAGCTACCGCCACAGCAGAGTAGCCGAGACTGGAGCC 1389

RESULT 9

US-09-178-115-1  
; Sequence 1, Application US/09178115  
; Patent No. 6297041  
; GENERAL INFORMATION:  
; APPLICANT: Zavada, Jan  
; APPLICANT: Pastorekova, Silvia  
; APPLICANT: Pastorek, Jaromir  
; TITLE OF INVENTION: MN Gene and Protein  
; FILE REFERENCE: D-0021.SA  
; CURRENT APPLICATION NUMBER: US/09/178,115  
; EARLIER FILING DATE: 1998-10-23  
; EARLIER APPLICATION NUMBER: 09/177,776  
; EARLIER FILING DATE: 1998-10-23  
; EARLIER APPLICATION NUMBER: 08/787,739  
; EARLIER FILING DATE: 1997-01-24  
; EARLIER APPLICATION NUMBER: 08/485,049  
; EARLIER FILING DATE: 1995-06-07  
; EARLIER APPLICATION NUMBER: 08/486,756  
; EARLIER FILING DATE: 1995-06-07  
; EARLIER APPLICATION NUMBER: 08/477,504  
; EARLIER FILING DATE: 1995-06-07  
; EARLIER APPLICATION NUMBER: 08/481,658  
; EARLIER FILING DATE: 1995-06-07  
; EARLIER APPLICATION NUMBER: 08/485,862  
; EARLIER FILING DATE: 1995-06-07  
; EARLIER APPLICATION NUMBER: 08/485,863  
; EARLIER FILING DATE: 1995-06-07  
; EARLIER APPLICATION NUMBER: 08/487,077  
; EARLIER FILING DATE: 1995-06-07  
; EARLIER APPLICATION NUMBER: 08/260,190  
; EARLIER FILING DATE: 1994-06-15  
; EARLIER APPLICATION NUMBER: 08/177,093  
; EARLIER FILING DATE: 1993-12-30  
; EARLIER APPLICATION NUMBER: 07/964,589  
; EARLIER FILING DATE: 1992-10-21  
; EARLIER APPLICATION NUMBER: PV-709-92  
; EARLIER FILING DATE: 1992-03-11  
; NUMBER OF SEQ ID NOS: 116  
; SOFTWARE: PatentIn ver. 2.0  
; SEQ ID NO 1  
; LENGTH: 1522  
; TYPE: DNA  
; ORGANISM: HUMAN  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (13)..(1389)  
; FEATURE:  
; NAME/KEY: mat\_peptide  
; LOCATION: (124)..(1389)  
US-09-178-115-1

Alignment Scores:  
Pred. No.: 7,5e-197 Length: 1522  
Score: 2424.00 Matches: 459  
Percent Similarity: 100.0% Conservative: 0  
Best Local Similarity: 100.0% Mismatches: 0  
Query Match: 100.0% Indels: 0  
DB: 3 Gaps: 0

US-09-967-237B-2 (1-459) x US-09-178-115-1 (1-1522)

Qy 1 MetAlaProLeuCysProSerProTyrLeuProLeuLeuIleProAlaProAlaProGly 20

Db 13 ATGGCTCCCTGTGCCCCAGCCCTGGCTCCCTCTGTGTGATCCCGGCCCTGTCTCAGGC 72  
Qy 21 LeuThrValGlnLeuLeuSerLeuLeuLeuMetProValHisProGlnHisArgLeu 40  
Db 73 CTCACCTGTGCAACTGTCTGTCTCCTGCTGCTCTCTGATGCCTGTCATCCACAGAGGTG 132  
Qy 41 ProArgMetGlnGluAspSerProLeuGlyGlySerSerGlyGluAspAspProLeu 60  
Db 133 CCCCAGATCAGAGAGATTCCTCCCTTTGGGAGGAGGCTCTTCTGGGAAAGATGATCCACTG 192  
Qy 61 GlyGluGluAspLeuProSerGluGluAspSerProArgGluGluAspProProGlyGlu 80  
Db 193 GGCAGAGAGATCTGCCAGTGAAGAGGATTCACCCAGAGAGAGAGATCCACCCGAGAG 252  
Qy 81 GluAspLeuProGlyGluGluAspLeuProGlyGluGluAspLeuProGluValIysPro 100  
Db 253 GAGGATCTACCTGGAGAGGAGATCTACCTGGAGAGGAGATCTACCTGAAGTTAAGCCT 312  
Qy 101 LysSerGluGluGlySerLeuLeuGluAspLeuProThrValGluAlaProGly 120  
Db 313 AAATCAGAAAGAGAGGCTCCCTGAAGTTAGAGGATCTACCTACTGTTGAGGCTCTCGGA 372  
Qy 121 AspProGlnGluProGlnAsnAsnAlaHisArgAspLysGluGlyAspAspGlnSerHis 140  
Db 373 GATCTCTCAAGAACCCAGAAATATGCCACAGGACAAAGAGGATGACACAGATCAT 432  
Qy 141 TrpArgTyrGlyGlyAspProProTyrProArgValSerProAlaCysAlaGlyArgPhe 160  
Db 433 TGGCGCTATGGAGGCGACCCGCTGGCGCCCGGGTGTCCCGAGCTGCGGGCGGCTTC 492  
Qy 161 GlnSerProValAspIleArgProGlnLeuAlaAlaPheCysProAlaLeuArgProLeu 180  
Db 493 CAGTCCCGCTGGATATCCCGCCCGAGCTCGCGGCTCTTCTGCCCGGCTCGCGCCCTCG 552  
Qy 181 GluLeuLeuGlyPheGlnLeuProProLeuProGluLeuArgLeuArgAsnGlyHis 200  
Db 553 GAACTCTGGGCTTCAGCTCCCGCGCTCCAGAACTCGCGCTCGCGCAACATATGGCCAC 612  
Qy 201 SerValGlnLeuThrLeuProGlyLeuGluMetAlaLeuGlyProGlyArgGluTyr 220  
Db 613 AGTGTGCAACTGACCTGCTCTCTGGCTTAGAGATGGCTCTGGGTCCCGGGGGAGTAC 672  
Qy 221 ArgAlaLeuGlnLeuHisLeuHisTrpGlyAlaAlaGlyArgProGlySerGluHisThr 240  
Db 673 CGGGCTCTGCAGCTGCATCTGCACCTGGGGGGTGCAGGTCGTCCGGGCTCGGAGCACAT 732  
Qy 241 ValGluGlyHisArgPheProAlaGluIleHisValHisValHisLeuSerThrAlaPheAla 260  
Db 733 GTGGAAGGCCACCGTTTCTCCCGAGATCCAGTGGTTACCTCAGCACCCGCTTTGCC 792  
Qy 261 ArgValAspGluAlaLeuGlyArgProGlyGlyLeuAlaValLeuAlaPheLeuGlu 280  
Db 793 AGAGTTGACGAGGCTTTGGGCGCCCGGAGGCTGGCGGTGTGGCGGCTTTCTGGAG 852  
Qy 281 GluGlyProGluGluAsnSerAlaTyrGluGlnLeuLeuSerArgLeuGluIleAla 300  
Db 853 GAGGGCCCGAAAGAAACAGTGCCTATGAGCAGTGTCTGCTGCTGTGGAGAAATCGCT 912  
Qy 301 GluGluGlySerGluThrGlnValProGlyLeuAspIleSerAlaLeuLeuProSerAsp 320  
Db 913 GAGGAAGGCTCAGAGACTCAGGTCCAGGACTGGACATATCTGACTCTCTGCCCTCTGAC 972  
Qy 321 PheSerArgTyrPheGlnTyrGluGlySerLeuThrThrProProCysAlaGlnGlyVal 340  
Db 973 TTCAGCCGCTACTTCCAATATGAGGGTCTCTGACTACACCGCCCTGTGCCAGGGTGC 1032  
Qy 341 IleTrpThrValPheAsnGlnThrValMetLeuSerAlaLysGlnLeuHisThrLeuSer 360  
Db 1033 ATCTGCACTGTGTTAAACCAGACAGTGTGCTGAGTGTGAAGCAGCTCCACCCCTCTCT 1092  
Qy 361 AspThrLeuTrpGlyProGlyAspSerArgLeuGlnLeuAsnPheArgAlaThrGlnPro 380

1093	GACACCCCTGGGGAGCTGCTGTGACTCTCGGCTACAGCTGAACCTTCGAGGCGACGAGCCT	1152
Db		
381	LeuAsnGlyArgValIleGluAlaSerPheProAlaGlyValAspSerSerProArgAla	400
QY		
1153	TTGAATGGGCGAGTGATTGAGGCTCTCTCCCTGCTGGAGTGGACAGCAGCTCCTCGGCT	1212
Db		
401	AlaGluProValGlnLeuAsnSerCysLeuAlaGlyAspIleLeuAlaLeuValPhe	420
QY		
1213	GCTGAGCCAGTCAGCTGAATTCCTGCTCGCTGCTGAGTGCATCCTAGGCCCTGGTFTTT	1272
Db		
421	GlyLeuLeuPheAlaValThrSerValAlaPheLeuValGlnMetArgArgGlnHisArg	440
QY		
1273	GGCCTCTCTTTTGTCTGTCACAGGTCGCGTCTCTGTGCGATGAGAGGCGACGACAGA	1332
Db		
441	ArgGlyThrLysGlyGlyValSerTyArgProAlaGluValAlaGluThrGlyAla	459
QY		
1333	AGGGGAACCAAGGGGTGTGTAGCTACGCCCCAGCAGAGTGGCCGAGATGGAGCC	1389
Db		

## RESULT 10

```

US-09-177-776-1
Sequence 1, Application US/09177776A
Patent No. 6297051
GENERAL INFORMATION:
APPLICANT: Zavada, Jan
APPLICANT: Pastorekova, Silvia
APPLICANT: Pastorek, Jaromir
TITLE OF INVENTION: MN Gene and Protein
FILE REFERENCE: D-0021.5A
CURRENT APPLICATION NUMBER: US/09/177,776A
CURRENT FILING DATE: 1998-10-23
EARLIER APPLICATION NUMBER: 08/787,739
EARLIER FILING DATE: 1997-01-24
EARLIER APPLICATION NUMBER: 08/485,049
EARLIER FILING DATE: 1995-06-07
EARLIER APPLICATION NUMBER: 08/486,756
EARLIER FILING DATE: 1995-06-07
EARLIER APPLICATION NUMBER: 08/477,504
EARLIER FILING DATE: 1995-06-07
EARLIER APPLICATION NUMBER: 08/481,658
EARLIER FILING DATE: 1995-06-07
EARLIER APPLICATION NUMBER: 08/485,862
EARLIER FILING DATE: 1995-06-07
EARLIER APPLICATION NUMBER: 08/485,863
EARLIER FILING DATE: 1995-06-07
EARLIER APPLICATION NUMBER: 08/487,077
EARLIER FILING DATE: 1995-06-07
EARLIER APPLICATION NUMBER: 08/260,190
EARLIER FILING DATE: 1994-06-15
EARLIER APPLICATION NUMBER: 08/177,093
EARLIER FILING DATE: 1993-12-30
EARLIER APPLICATION NUMBER: 07/964,589
EARLIER FILING DATE: 1992-10-21
EARLIER APPLICATION NUMBER: PW-709-92
EARLIER FILING DATE: 1992-03-11
NUMBER OF SEQ ID NOS: 116
SOFTWARE: Patentin Ver. 2.0
SEQ ID NO 1
LENGTH: 1522
TYPE: DNA
ORGANISM: HUMAN
FEATURE:
NAME/KEY: CDS
LOCATION: (13)..(1389)
FEATURE:
NAME/KEY: mat_peptide
LOCATION: (124)..(1389)
US-09-177-776-1

```

Alignment Scores:		
Pred. No.:	7.5e-197	1522
Score:	2424.00	459
Percent Similarity:	100.0%	Conservative: 0
Best Local Similarity:	100.0%	Mismatches: 0

Query Match:	100.0%	Indels:	0
DB:	3	Gaps:	0
US-09-967-237B-2 (1-459) x US-09-177-776-1 (1-1522)			
Qy	1	MetAlaProLeuCyvProSerProTtpLeuProLeuLeuIleProAlaProAlaProGly	20
Db	13	ATGGCTCCCTGFGCCCAAGCCCTGGCTCCCTCTGTGTATCCCGGCCCTGCTCCAGGC	72
Qy	21	LeuThrValGlnLeuLeuSerLeuLeuLeuMetProValHisProGlnArgLeu	40
Db	73	CTCACTGTGCAACTGCTGTCTCACTGCTCTCTGATGCTGTCTCCATCCCCAGAGTTG	132
Qy	41	ProArgMetGlnGluAspSerProLeuGlyGlyGlySerSerGlyGluAspAspProLeu	60
Db	133	CCCCGGATGCAGAGGATTCCTCCCTTGGGAGGAGGCTCTTCTGGGGAAGATGACCCACTG	192
Qy	61	GlyGluGluAspLeuProSerGluGluAspSerProArgGluGluAspProProGlyGlu	80
Db	193	GGCGAGAGGATCTGCCAGTGAAGAGATTACCCAGAGAGAGGATCCACCCGGAGAG	252
Qy	81	GluAspLeuProGlyGluGluAspLeuProGlyGluGluAspLeuProGluValHisPro	100
Db	253	GAGGATCTACCTGGAGAGGAGTCTACCTGGAGAGGAGGATCTACCTGAAGTTAAGCCT	312
Qy	101	LysSerGluGluGlySerLeuHisLeuGluAspLeuProThrValGluAlaProGly	120
Db	313	AAATCAAGAAAGAGGCTCCCTCGAAGTTAGAGGATCTACCTACTGTTGAGGCTCTCGGA	372
Qy	121	AspProGlnGluProGlnAenAenAlaHisArgAspLysGluGlyAspAspGlnSerHis	140
Db	373	GATCTCTCAAGAACCCCAAGATATATGCCACAGGACAAAGAGGGATGACACAGATCAT	432
Qy	141	TtpArgTyrGlyGlyAspProTtpProArgValSerProAlaCysAlaGlyArgPhe	160
Db	433	TGGCGCTATGGAGGCGACCCGCGCTGCGCCCGGGTGTCCACAGCTGCGCGGCGCTTC	492
Qy	161	GlnSerProValAspIleArgProGlnLeuAlaAlaPheCysProAlaLeuArgProLeu	180
Db	493	CAGTCCCGGTGATATCCGCCCCAGCTCGCCGCTTCTGCCCGGCCCTTGCGCCCCCTG	552
Qy	181	GluLeuLeuGlyPheGlnLeuProLeuProGluLeuArgLeuArgAenAenGlyHis	200
Db	553	GAATCTCTGGGCTTCAGCTCCCGCGCTCCCAAGACTGGCGCTGGCGACAATGGCCAC	612
Qy	201	SerValGlnLeuThrLeuProProGlyLeuGluMetAlaLeuGlyProGlyArgGluTyr	220
Db	613	AGTGTGCAACTGACCCCTGCTCTCTGGGCTAGAGATGGCTCTGGGTCGCCGGCGGGAGTAC	672
Qy	221	ArgAlaLeuGlnLeuHisLeuHisTtpGlyValAlaAlaGlyArgProGlySerGluHisThr	240
Db	673	CGGGCTCTGCAGTGCATCTGCAGTGGGGGGCTGCGAGTGTCTGGGCTCGAGACACACT	732
Qy	241	ValGluGlyHisArgPheProAlaGluIleHisValHisLeuSerThrAlaPheAla	260
Db	733	GTGGAAGGCCACCGTTCCTTCCGAGATCCAGTGGTTCACCTCAGCACCGGCTTTGCC	792
Qy	261	ArgValAspGluAlaLeuGlyArgProGlyGlyLeuAlaValLeuAlaAlaPheLeuGlu	280
Db	793	AGAGTTGACGAGGCTTGGGGCGCCCGGAGGCTGGCCGTGTGTGGCGGCTCTTCTGGAG	852
Qy	281	GluGlyProGluGluAenSerAlaTyrGluGlnLeuLeuSerArgLeuGluGluIleAla	300
Db	853	GAGGGCCCGGAAGAAACAGTGGCTATGAGCAGTTGCTGTCTCGTTGGGAAGAAATCGCT	912
Qy	301	GluGluGlySerGluThrGlnValProGlyLeuAspIleSerAlaLeuLeuProSerAsp	320
Db	913	GAGGAAGGCTCAGAGACTCAGTCCAGTCCAGGACTGGACATATCTGCATCTCTGCTCTGAC	972
Qy	321	PheSerArgTyrPheGlnTyrGluGlySerLeuThrThrProProCysAlaGlnGlyVal	340
Db	973	TTCAAGCCGCTACTTCCAAATATGAGGGTCTCTGACTACACGGCTCTGCGCCAGGGTGC	1032

QY 341 IletpThrValPheAenGlnThrValMetLeuSerAlaLysGlnLeuHisThrLeuSer 360  
Db 1033 ATCTGGAGCTGTGTTTACACAGACAGTAGTCTGAGTCTAAGCAGCTCCACACCCCTCTCT 1092  
QY 361 AspThrLeuTrpGlyProGlyAspSerArgLeuGlnLeuAenPheArgAlaThrGlnPro 380  
Db 1093 GACACCCCTGTGGGAGCTGTGAGCTCTCGGTACAGCTGAACCTCCGAGCGACGACGCT 1152  
QY 381 LeuAenGlyArgValIleGluAlaSerPheProAlaGlyValAspSerSerProArgAla 400  
Db 1153 TTGAATGGGAGTAGTATTGAGGCTCTCTCCCTGCTGAGTAGGACAGCAGCTCTCGGGCT 1212  
QY 401 AlaGluProValGlnLeuAenSerCysLeuAlaGlyAspIleLeuAlaLeuValPhe 420  
Db 1213 GCTGAGCCAGTCACGCTGAATCTCTGCTGGCTGCTGCTGACATCTTAGCCCTGGTGT 1272  
QY 421 GlyLeuLeuPheAlaValThrSerValAlaPheLeuValGlnMetArgArgGlnHisArg 440  
Db 1273 GGCTCTCTTTTGTCTGTCACCGCTGCGCTTCTTGTGAGATGAGAGCGACACAGA 1332  
QY 441 ArgGlyThrLysGlyValSerTyrArgProAlaGluValAlaGluThrGlyAla 459  
Db 1333 AGGGGAACCAAGGGGTGTGAGCTACCGCCAGCAGAGGTAGCCGAGACTGGAGCC 1389

## RESULT 11

US-09-772-719B-1  
; Sequence 1, Application US/09772719B  
; Patent No. 6770438  
; GENERAL INFORMATION:  
; APPLICANT: Zavada, Jan  
; Pastorekova, Silvia  
; Pastorek, Jaromir  
; TITLE OF INVENTION: MN Gene and Protein  
; NUMBER OF SEQUENCES: 86  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Leona L. Lauder  
; STREET: 465 California Street, Suite 450  
; CITY: San Francisco  
; STATE: California  
; COUNTRY: USA  
; ZIP: 94104  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent in Release #1.0, Version #1.30 (EPO)  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/772,719B  
; FILING DATE: 30-Jan-2001  
; CLASSIFICATION: <Unknown>  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/485,049  
; FILING DATE: 07-JUN-1995  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Lauder, Leona L.  
; REGISTRATION NUMBER: 30,863  
; REFERENCE/DOCKET NUMBER: D-0021.3A-2  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 415-981-2034  
; TELEFAX: 415-981-0332  
; INFORMATION FOR SEQ ID NO: 1:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1522 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: cDNA  
; HYPOTHETICAL: NO  
; ANTI-SENSE: NO  
; SEQUENCE DESCRIPTION: SEQ ID NO: 1:  
US-09-772-719B-1  
Alignment Scores:

Pred. No.: 7 Se-197 Length: 1522  
Score: 2424.00 Matches: 459  
Percent Similarity: 100.0% Conservatives: 0  
Best Local Similarity: 100.0% Mismatches: 0  
Query Match: 100.0% Indels: 0  
DB: 3 Gaps: 0  
US-09-967-237B-2 (1-459) x US-09-772-719B-1 (1-1522)  
QY 1 MetAlaProLeuCysProSerProTrpLeuProLeuLeuLeuProAlaProGly 20  
Db 13 ATGGCTCCCTGTGCCCCAGCCCTGGCTCCCTCTGTGATCCCGCCCTGTCTCCAGGC 72  
QY 21 LeuThrValGlnLeuLeuSerLeuLeuLeuMetProValHisProGlnArgLeu 40  
Db 73 CTCACCTGTGCAACTGTGCTGTCTGCTGCTCTGATGCTCTGATGCTCTCCACAGAGGTTG 132  
QY 41 ProArgMetGlnGluAspSerProLeuGlyGlyGlySerSerGlyGluAspProLeu 60  
Db 133 CCCCCGATGAGGAGGATTCCTCCCTTGGAGAGGAGGCTCTTCTGGGGAAGATGATCCCACTG 192  
QY 61 GlyGluGluAspLeuProSerGluGluAspSerProArgGluGluAspProGlyGlu 80  
Db 193 GCGAGAGGAGTCTGCCAGTGAAGAGGATTCACCCAGAGAGGAGATCCACCCGAGAG 252  
QY 81 GluAspLeuProGlyGluGluAspLeuProGlyGluGluAspLeuProGluValLysPro 100  
Db 253 GAGGATCTACCTGGAGAGGAGATCTACCTGGAGAGGAGGATCTACCTGAAGTTAGCCT 312  
QY 101 LysSerGluGluGlySerLeuLysLeuGluAspLeuProThrValGluAlaProGly 120  
Db 313 AAATCAGAAGAGAGGGCTCCCTGAAAGTTAGAGGATCTACCTACTGTTGAGGCTCTCGGA 372  
QY 121 AspProGlnGluProGlnAsnAlaHisArgAspLysGluGlyAspAspGlnSerHis 140  
Db 373 GATCTCTCAAGAACCCCGAGAAATGATGCCACAGGACAAAGAGGGGATGACACAGTCTAT 432  
QY 141 TrpArgTyrGlyGlyAspProProTrpProArgValSerProAlaCysAlaGlyArgPhe 160  
Db 433 TGGCGCTATGAGAGGCGACCGCCCTGGCGCCCGGGTGTCCCGAGCTTCGCGGCGCGCTTC 492  
QY 161 GlnSerProValAspIleArgProGlnLeuAlaAlaPheCysProAlaLeuArgProLeu 180  
Db 493 CAGTCCCCGGTGGATATCCGCCCGCCAGCTCGCGCCCTTCTGCCGCGCTTGGCCCCCTTG 552  
QY 181 GluLeuLeuGlyPheGlnLeuProProLeuProGluLeuArgLeuArgAsnGlnHis 200  
Db 553 GAACCTCTGGCTTCCAGCTCCCGCGCTCCCGAGAACTGCGCTGCGCAACAATGGCCAC 612  
QY 201 SerValGlnLeuThrLeuProProGlyLeuGluMetAlaLeuGlyProGlyArgGluTyr 220  
Db 613 AGTGTGCAACTGACCTGCTCTCTGGGCTAGAGATGGCTCTGGGTCCCGGGCGGAGTAC 672  
QY 221 ArgAlaLeuGlnLeuHisLeuHisTrpGlyAlaAlaGlyArgProGlySerGluHisThr 240  
Db 673 CGGGCTCTGAGCTGCTCTGCTGCGGGGGGTGACAGGTCTGTCGCGGCTCGGAGCACACT 732  
QY 241 ValGluGlyHisArgPheProAlaGluIleHisValHisValHisLeuSerThrAlaPheAla 260  
Db 733 GTGGAAGGCCACCGCTTTCCTGCGAGATCCACGTGTTTCACTCAGCACCGCGCTTTTGCC 792  
QY 261 ArgValAspGluAlaLeuGlyArgProGlyGlyLeuAlaValLeuAlaAlaPheLeuGlu 280  
Db 793 AGAGTTGACGAGGCTTGGGGCGCGCGGAGGCGCTGGCGCTGTGGCCCTTTCTGGAG 852  
QY 281 GluGlyProGluGluAsnSerAlaTyrGluGlnLeuLeuSerArgLeuGluLeuAla 300  
Db 853 GAGGGCCCGAAGAAACAGTGCCTATGAGCAGTTGCTCTCTCGCTTGAAGAAATCGCT 912  
QY 301 GluGluGlySerGluThrGlnValProGlyLeuAspIleSerAlaLeuLeuProSerAsp 320  
Db 913 GAGGAAGGCTCAGAGACTCAGGTCCCGAGACTGGACATATCTGCATCTCTGCTCTGAC 972

```
QY 321 PheSerArgTyrPheGlnTyrGluGlySerLeuThrThrProCysAlaGlnGlyVal 340
Db 973 TTGAGCGGTACTTCCAAATAGAGGGTCTCTGACTACACCGCCCTGTGCCAGGGTGTG 1032
QY 341 IleTrpThrValPheAsnGlnThrValMetLeuSerAlaLysGlnLeuHisThrLeuSer 360
Db 1033 ATCTGACTGTGTTTAAACACAGACAGTGTCTGAGTGTCTAAGCAGCTCCACACCCCTCTCT 1092
QY 361 AspThrLeuTrpGlyProGlyAspSerArgLeuGlnLeuAsnPheArgAlaThrGlnPro 380
Db 1093 GACACCCCTGTGGGACCTGTGTGACTCTCGGTACTACAGTGAACCTTCGAGCGACGAGCCT 1152
QY 381 LeuAsnGlyArgValIleGluAlaSerPheProAlaGlyValAspSerSerProArgAla 400
Db 1153 TTGAATGGGCGAGTGTAGGGCTCTCTCCCTGCTGGAGTGGACAGCAGTCTCTCGGGCT 1212
QY 401 AlaGluProValGlnLeuAsnSerCysLeuAlaAlaGlyAspIleLeuAlaLeuValPhe 420
Db 1213 GCTGAGCCAGTCCAGCTGAATTCCTGCTGCTGCTGGTGACATCTAGCCCTGTGTTTTT 1272
QY 421 GlyLeuLeuPheAlaValThrSerValAlaPheLeuValGlnMetArgArgGlnHisArg 440
Db 1273 GGCCCTCTTTTGTCTGTACACAGCTGCGGTTCCTTTGTGCAGATGAGAGGCAGCACAGA 1332
QY 441 ArgGlyThrLysGlyValSerTyrArgProAlaGluValAlaGluThrGlyAla 459
Db 1333 AGGGGAACCAAGAGGGGTGTGAGCTACCGCCACGAGAGTAGCCGAGACTGGAGCC 1389
```

## RESULT 12

```
US-08-260-190-5
; Sequence 5, Application US/08260190A
; Patent No. 6774117
; GENERAL INFORMATION:
; APPLICANT: Zavada, Jan
; APPLICANT: Pastorekova, Silvia
; APPLICANT: Pastorek, Jaromir
; TITLE OF INVENTION: MN Gene and Protein
; FILE REFERENCE: D-0021-2
; CURRENT APPLICATION NUMBER: US/08/260,190A
; CURRENT FILING DATE: 1994-06-15
; EARLIER APPLICATION NUMBER: 08/177,093
; EARLIER FILING DATE: 1993-12-30
; EARLIER APPLICATION NUMBER: 07/964,589
; EARLIER FILING DATE: 1992-10-21
; EARLIER APPLICATION NUMBER: PV-709-92
; EARLIER FILING DATE: 1992-03-11
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 5
; LENGTH: 1522
; TYPE: DNA
; ORGANISM: HUMAN
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (13)..(1389)
; FEATURE:
; NAME/KEY: mat_peptide
; LOCATION: (124)..(1389)
US-08-260-190-5
```

```
Alignment Scores:
Pred. No.: 7.5e-197 Length: 1522
Score: 2424.00 Matches: 459
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 3 Gaps: 0
```

US-09-967-237B-2 (1-459) x US-08-260-190-5 (1-1522)

```
QY 1 MetAlaProLeuCysProSerProTrpLeuProLeuLeuIleProAlaProGly 20
Db 13 ATGGCTCCCTGTGCCCCAGCCCTGGCTCCCTCTGTGTGATCCGGGCCCTGTCTCCAGGC 72
```

```
QY 21 LeuThrValGlnLeuLeuSerLeuLeuLeuMetProValHisProGlnArgLeu 40
Db 73 CTCACCTGTGCACACTGCTGCTGTCTGCTGCTCTTCTGATGCTCTCATCCACAGAGGTG 132
QY 41 ProArgMetGlnGluAspSerProLeuGlyGlySerSerGlyGluAspAspProLeu 60
Db 133 CCCCGGATGCAGGAGGATTTCCCTTGGGAGGAGGCTCTTCTGGGGAAGATGACCCACTG 192
QY 61 GlyGluGluAspLeuProSerGluGluAspSerProArgGluGluAspProProGlyGlu 80
Db 193 GCGAGAGAGGATCTGCCAGTAGAGGAGGATTCACCCAGAGAGAGGATTCACCCGAGAG 252
QY 81 GluAspLeuProGlyGluGluAspLeuProGlyGluGluAspLeuProGluValLysPro 100
Db 253 GAGGATCTACTTGGAGAGGAGGATCTACTCTGGAGAGGAGGATCTACTTGAAGTTAAGCCT 312
QY 101 LysSerGluGluGlySerLeuLysLeuGluAspLeuProThrValGluAlaProGly 120
Db 313 AAATCAGAAAGAGAGGCTCCCTGAAGTTAGAGGATCTACTTCTGTGAGGCTCTCTGGA 372
QY 121 AspProGlnGluProGlnAsnAsnAlaHisArgAspLysGluGlyAspAspGlnSerHis 140
Db 373 GATCTCTCAAGAACCCAGAAATAATGCCACAGGACAAAGAGGAGGATGACACAGAGTCAT 432
QY 141 TrpArgTyrGlyGlyAspProProTrpProArgValSerProAlaCysAlaGlyArgPhe 160
Db 433 TGGCGGTATGGAGGAGGACCCGCGCTGGCCCGGGTGTCTCCACAGCTTCGCGGSCGCTTC 492
QY 161 GlnSerProValAspIleArgProGlnLeuAlaAlaPheCysProAlaLeuArgProLeu 180
Db 493 CAGTCCCGGTGGATATCCGCGCCACAGCTCGCGCGCTTCTGCGCGGCTTCGCGCCCTCG 552
QY 181 GluLeuLeuGlyPheGlnLeuProProLeuProGluLeuArgLeuArgAsnGlyHis 200
Db 553 GAACTCTGGGCTTCCAGCTCCGCGCTCCAGAACTCGCGCTCGCGCAACATGGCCAC 612
QY 201 SerValGlnLeuThrLeuProProGlyLeuGluMetAlaLeuGlyProGlyArgGluTyr 220
Db 613 AGTGTGCACACTGACCTTGCTCTCTGGCTAGAGATGGCTCTGGTCCCGCGGGAGTAC 672
QY 221 ArgAlaLeuGlnLeuHisLeuHisTrpGlyAlaAlaGlyArgProGlySerGluHisThr 240
Db 673 CCGGCTCTGCAGCTGCATCTGCACCTGGGGGCTGCAGGTCGTCCGGGCTCGAGACACACT 732
QY 241 ValGluGlyHisArgPheProAlaGluIleHisValHisLeuSerThrAlaPheAla 260
Db 733 GTGGAAGGCCACCGTTTCCCTGCGGAGATCCACGTGGTTTCCCTCAGCACCGCTTTGGCC 792
QY 261 ArgValAspGluAlaLeuGlyArgProGlyGlyLeuAlaValLeuAlaAlaPheLeuGlu 280
Db 793 AGAGTTGACAGGCGCTTGGGGCGCCCGGGAGGCGCTGGCGGTGTGGCGGCTTTCTGGAG 852
QY 281 GluGlyProGluGluAsnSerAlaTyrGluGlnLeuLeuSerArgLeuGluIleAla 300
Db 853 GAGGGCCCCGAGAGAAACAGTGCCTATGAGCAGTTGCTGTCTGCTTGGAAAGAAATCGCT 912
QY 301 GluGluGlySerGluThrGlnValProGlyLeuAspIleSerAlaLeuLeuProSerAsp 320
Db 913 GAGGAAGGCTCAGAGACTCAGGTCCAGGAGACTGGACATATCTGCACATCTCCCTCTGAC 972
QY 321 PheSerArgTyrPheGlnTyrGluGlySerLeuThrThrProCysAlaGlnGlyVal 340
Db 973 TTCAGCGGTACTTCCAAATAGAGGGTCTCTGACTACACCGCCCTGTGCCAGGGTGTG 1032
QY 341 IleTrpThrValPheAsnGlnThrValMetLeuSerAlaLysGlnLeuHisThrLeuSer 360
Db 1033 ATCTGACTGTGTTTAAACACAGACAGTGTCTGAGTGTCTAAGCAGCTCCACACCCCTCTCT 1092
QY 361 AspThrLeuTrpGlyProGlyAspSerArgLeuGlnLeuAsnPheArgAlaThrGlnPro 380
Db 1093 GACACCCCTGTGGGAGACTGTGTGACTCTCTGGCTACAGCTGAACCTTCGAGCGACGAGCCT 1152
```



Db	643	AGTGTGCAACTGACCTGCTCTCTGGGCTAGAGTGGCTCTGGGTCCCGGCGGGAGTAC	702
Qy	221	ArgAlaLeuGlnLeuHisLeuHisTrrpGlyAlaAlaGlyArgProGlySerGluHisThr	240
Db	703	CGGGCTCTCGAGCTGCATCTGCACATCGGGGGCTGCAGGTGCTCGGGCTCGGAGCACACT	762
Qy	241	ValGluGlyHisArgPheProAlaGluLleHisValValHisLeuSerThrAlaPheAla	260
Db	763	GTGGAAAGGCCACCGTTTCCTTGCAGAGATCCACGTGGTTCCACTCAGCACCGGCTTTGCC	822
Qy	261	ArgValAspGluAlaLeuGlyArgProGlyGlyLeuAlaValLeuAlaAlaPheLeuGlu	280
Db	823	AGAGTTGACAGAGCCCTTGGGGCGCCCGGAGGCTCGCGTGTGTGGCCGCTTCTTGGAG	882
Qy	281	GluGlyProGluGluAsnSerAlaTyrGluGlnLeuLeuSerArgLeuGluLleAla	300
Db	883	GAGGGCCCGGAAGAAACAGTGCCTATGAGCAGTCTGTCTCGTTGGAAGAAATCGCT	942
Qy	301	GluGluGlySerGluThrGlnValProGlyLeuAspIleSerAlaLeuLeuProSerAsp	320
Db	943	GAGGAAGGCTCAGAGACTCAGGTCCCAGGACTGGACATATCTGCACCTCTGCCCTCTGAC	1002
Qy	321	PheSerArgTyrPheGlnTyrGluGlySerLeuThrProProCysAlaGlnGlyVal	340
Db	1003	TTCAGCCCGCTACTTCCAAATATGAGGGGTCTCTGACTACACCGCCCTGTGCCAGGGGTGC	1062
Qy	341	IleTrrpThrValPheAsnGlnThrValMetLeuSerAlaLysGlnLeuHisThrLeuSer	360
Db	1063	ATCTGAGACTGTGTTTAAACAGACAGTGTGCTGAGTGTCTAAGCAGCTCCACACCCCTCTCT	1122
Qy	361	AspThrLeuTrrpGlyProGlyAspSerArgLeuGlnLeuAsnPheArgAlaThrGlnPro	380
Db	1123	GACACCCCTGTGGGACCTGGTGACTCTCGGCTACAGCTGAACCTTCGAGCGACGACGCT	1182
Qy	381	LeuAsnGlyArgValIleGluAlaSerPheProAlaGlyValAspSerSerProArgAla	400
Db	1183	TTGAATGGGCGAGTGATTCAGAGCCCTCTTCCCTGCTGGAGTGGACAGCAGTCTCGGGCT	1242
Qy	401	AlaGluProValGlnLeuAsnSerCysLeuAlaAlaGlyAspIleLeuAlaLeuValPhe	420
Db	1243	GCTGAGCCAGTCAGCTGATTCCTGCCCTGGCTGTGGTGACATCTCTAGCCCTGGTTTTT	1302
Qy	421	GlyLeuLeuPheAlaValThrSerValAlaPheLeuValGlnMetArgArgGlnHisArg	440
Db	1303	GGCCTCTCTTTTGTCTGTCCAGCGTGGCGTTCCTTGTGCAGATGAGAAGGACAGCACAGA	1362
Qy	441	ArgGlyThrLysGlyValSerTyrArgProAlaGluValAlaGlnThrGlyAla	459
Db	1363	AGGGGAACCAAAAGGGGGTGTGAGTACCGCCCGACAGAGGTAGCCGAGACTGGAGCC	1419

RESULT 15

US-08-260-190-23

; Sequence 23, Application US/08260190A

; Patent No. 6774117

; GENERAL INFORMATION:

; APPLICANT: Zavada, Jan

; APPLICANT: Pastorekova, Silvia

; APPLICANT: Pastorek, Jaromir

; TITLE OF INVENTION: MN Gene and Protein

; FILE OF INVENTION: D-0021-2

; CURRENT APPLICATION NUMBER: US/08/260,190A

; CURRENT FILING DATE: 1994-06-15

; EARLIER APPLICATION NUMBER: 08/177,093

; EARLIER FILING DATE: 1993-12-30

; EARLIER APPLICATION NUMBER: 07/964,589

; EARLIER FILING DATE: 1992-10-21

; EARLIER APPLICATION NUMBER: PV-709-92

; EARLIER FILING DATE: 1992-03-11

; NUMBER OF SEQ ID NOS: 26

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 23

; LENGTH: 5052



GenCore version 5.1.7  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: February 17, 2006, 15:26:00 ; Search time 1362 Seconds  
(without alignments)  
2786.817 Million cell updates/sec

Title: US-09-967-237B-2

Perfect score: 2424

Sequence: 1 MAPLCSPWMLPLIPAPAG.....RRGTGGVSYPRAEVAETGA 459

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 9793542 seqs, 4134689005 residues

Total number of hits satisfying chosen parameters: 19587084

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlip  
-Q=/abs/ABSSWEB.spool/US0967237/runat\_15022006.182254.28230/app\_query.fasta\_1  
-DB=Published Applications NA Main -QPMT=fastap -SUFFIX=rnpbm -MINMATCH=0.1  
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blos62  
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100  
-THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0  
-MAXLEN=2000000000 -HOST=abs06p  
-USER=US0967237 @CGN 1.1.1549 @runat\_15022006.182254.28230 -NCPU=6 -ICPU=3  
-NO MMAP -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG -DEV TIMEOUT=120  
-WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7  
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications NA Main:

1: /cgn2\_6/ptodata/1/pubpna/US07\_PUBCOMB.seq:\*  
2: /cgn2\_6/ptodata/1/pubpna/US08\_PUBCOMB.seq:\*  
3: /cgn2\_6/ptodata/1/pubpna/US09A\_PUBCOMB.seq:\*  
4: /cgn2\_6/ptodata/1/pubpna/US09B\_PUBCOMB.seq:\*  
5: /cgn2\_6/ptodata/1/pubpna/US10A\_PUBCOMB.seq:\*  
6: /cgn2\_6/ptodata/1/pubpna/US10B\_PUBCOMB.seq:\*  
7: /cgn2\_6/ptodata/1/pubpna/US10C\_PUBCOMB.seq:\*  
8: /cgn2\_6/ptodata/1/pubpna/US10D\_PUBCOMB.seq:\*  
9: /cgn2\_6/ptodata/1/pubpna/US10E\_PUBCOMB.seq:\*  
10: /cgn2\_6/ptodata/1/pubpna/US11\_PUBCOMB.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2424	100.0	1519	5	US-10-102-524-1695
2	2424	100.0	1522	3	Sequence 1695, Ap
3	2424	100.0	1522	3	Sequence 1, Appli
4	2424	100.0	1522	7	US-09-967-237-1
5	2424	100.0	1522	8	US-10-723-795-1
6	2424	100.0	1522	8	US-10-795-933-5
7	2424	100.0	1522	8	US-10-888-694-1
	2424	100.0	1522	8	US-10-921-590-1

8	2424	100.0	1552	3	US-09-954-456-89	Sequence 89, Appli
9	2424	100.0	1552	3	US-09-954-456-726	Sequence 726, App
10	2424	100.0	1552	3	US-09-960-706-1080	Sequence 1080, Ap
11	2424	100.0	1552	3	US-09-873-367C-516	Sequence 516, App
12	2424	100.0	1552	3	US-09-968-007A-213	Sequence 213, App
13	2424	100.0	1552	6	US-10-301-822-11	Sequence 11, Appli
14	2424	100.0	1552	6	US-10-465-572-9	Sequence 9, Appli
15	2424	100.0	1552	6	US-10-172-118-574	Sequence 574, App
16	2424	100.0	1552	6	US-10-388-360-291	Sequence 291, App
17	2424	100.0	1552	6	US-10-295-027-305	Sequence 305, App
18	2424	100.0	1552	6	US-10-295-027-1022	Sequence 1022, Ap
19	2424	100.0	1552	7	US-10-342-887-574	Sequence 574, App
20	2424	100.0	1552	7	US-10-734-564-71	Sequence 71, Appli
21	2424	100.0	1552	8	US-10-723-860-460	Sequence 460, App
22	2424	100.0	1552	8	US-10-921-590-70	Sequence 70, Appli
23	2424	100.0	1552	9	US-10-848-755A-94	Sequence 94, Appli
24	2424	100.0	1552	9	US-10-843-641A-516	Sequence 516, App
25	2424	100.0	1552	9	US-10-843-641A-3116	Sequence 3116, App
26	2424	100.0	1552	9	US-10-843-641A-3753	Sequence 3753, Ap
27	2424	100.0	1552	9	US-10-843-641A-6683	Sequence 6683, Ap
28	2424	100.0	1552	9	US-10-794-514A-295	Sequence 295, App
29	2424	100.0	1552	9	US-10-756-149-474	Sequence 474, App
30	2424	100.0	1552	8	US-10-723-860-5135	Sequence 5135, Ap
31	2419	99.8	1833	3	US-09-783-708-2	Sequence 2, Appli
32	2416	99.7	1572	9	US-10-936-626-14	Sequence 14, Appli
33	2416	99.7	1572	9	US-10-938-061-14	Sequence 14, Appli
34	2233	92.1	5052	8	US-10-795-933-23	Sequence 23, Appli
35	2227	91.9	1399	8	US-10-795-933-1	Sequence 1, Appli
36	2058	84.9	1248	9	US-10-794-514A-321	Sequence 321, App
37	1938	80.0	1215	9	US-10-794-514A-323	Sequence 323, App
38	1596	65.8	1965	8	US-10-921-590-79	Sequence 79, Appli
39	1176.5	48.5	10898	3	US-09-772-719-5	Sequence 5, Appli
40	1176.5	48.5	10898	3	US-09-967-237-5	Sequence 5, Appli
41	1176.5	48.5	10898	7	US-10-723-795-3	Sequence 3, Appli
42	1176.5	48.5	10898	8	US-10-888-694-5	Sequence 5, Appli
43	1176.5	48.5	10898	8	US-10-921-590-3	Sequence 3, Appli
44	988	40.8	586	8	US-10-921-590-73	Sequence 73, Appli
45	714	29.5	415	3	US-09-772-719-28	Sequence 28, Appli

#### ALIGNMENTS

##### RESULT 1

US-10-102-524-1695  
; Sequence 1695, Application US/10102524  
; Publication No. US20030109434A1  
; GENERAL INFORMATION:  
; APPLICANT: Algate, Paul A.  
; APPLICANT: Mannion, Jane  
; APPLICANT: Gaiger, Alexander  
; APPLICANT: Gordon, Brian  
; APPLICANT: Harlocker, Susan L.  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND DIAGNOSIS OF KIDNEY CANCER  
; FILE REFERENCE: 210121.572  
; CURRENT APPLICATION NUMBER: US/10/102,524  
; CURRENT FILING DATE: 2002-03-19  
; NUMBER OF SEQ ID NOS: 1863  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 1695  
; LENGTH: 1519  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-10-102-524-1695

Alignment Scores:  
Pred. No.: 1.3e-227  
Score: 2424.00  
Matches: 1519  
Percent Similarity: 100.0%  
Best Local Similarity: 100.0%  
Query Match: 100.0%  
DB: 5  
Gaps: 0  
Indels: 0  
Mismatch: 0  
Conservative: 0  
Matches: 459  
Length: 1519

US-09-967-237B-2 (1-459) x US-10-102-524-1695 (1-1519)

QY 1 MetAlaProLeuCysProSerProTTPLeuProLeuLeuIleProAlaProAlaProGly 20  
Db 10 ATGGCTCCCTGTGCCCCAGCCCTGGCTCCCTCTGTGTGATCCGGCCCTGTCTCCAGGC 69  
QY 21 LeuThrValGlnLeuLeuSerLeuLeuLeuMetProValHisProGlnArgLeu 40  
Db 70 CTCACGTGTCAACTGCTGTCTCACTGCTGCTTCTGATGCTGTCCATCCCCAGAGGTTG 129  
QY 41 ProArgMetGlnGluAspSerProLeuGlyGlySerSerGlyGluAspAspProLeu 60  
Db 130 CCCCGGATGACAGAGGATCCCTTGGGAGGAGGCTTCTTGGGGAAGATGACCCACTG 189  
QY 61 GlyGluGluAspLeuProSerGluGluAspSerProArgGluGluAspProProGlyGlu 80  
Db 190 GCGGAGGAGGATCTGCCAGTGAAGAGGATTCACCCAGAGAGGAGGATTCACCCGGAGAG 249  
QY 81 GluAspLeuProGlyGluGluAspLeuProGlyGluGluAspLeuProGluValHisPro 100  
Db 250 GAGGATCTACCTGGAGAGGAGGATCTACCTGGAGAGGAGGATCTACCTGAAGTTAAGCCT 309  
QY 101 LysSerGluGluGlySerLeuIleLeuLeuAspLeuProThrValGluAlaProGly 120  
Db 310 AAATCAGAAGAAGAGGCTCCCTGAAGTTAGAGATCTACTACTGTGTAGGGCTCTCTGGA 369  
QY 121 AspProGlnGluProGlnAsnAlaHisArgAspLysGluGlyAspAspGlnSerHis 140  
Db 370 GATCCTCAAGAACCCAGATATATGCCACAGGACAAAGAGGAGTACACAGAGTCAT 429  
QY 141 TrpArgTyrGlyGlyAspProProTTPProArgValSerProAlaCysAlaGlyArgPhe 160  
Db 430 TGGCGCTATGAGGCGCACCCGCTTGGCCCGGGGTGTCCAGCTGCGCGGGCGGCTTC 489  
QY 161 GlnSerProValAspIleArgProGlnLeuAlaPheCysProAlaLeuArgProLeu 180  
Db 490 CAGTCCCGGTGATATCCCGCCCCAGCTCGCCGCTTCTGCCCCGGCCCTGCGCCCCCTG 549  
QY 181 GluLeuLeuGlyPheGlnLeuProProLeuProGluLeuArgLeuArgAsnAnglyHis 200  
Db 550 GAATCCTGGGCTTCCAGCTCCCGCCGCTCCAGAACTGGCGCTCGGCAACATGGCCAC 609  
QY 201 SerValGlnLeuThrLeuProProGlyLeuGluMetAlaLeuGlyProGlyArgGluTyr 220  
Db 610 AGTGTGCAACTGACCTGCTCTGGCTAGAGATGGCTCTGGGTCCCGGGCGGAGTAC 669  
QY 221 ArgAlaLeuGlnLeuHisLeuHisTTPGlyValAlaGlyArgProGlySerGluHisThr 240  
Db 670 CCGGCTCTGCAGCTGCATCTGCACCTGGGGGGCTGCAGGTGCTCGGGCTCGGAGCACAT 729  
QY 241 ValGluGlyHisArgPheProAlaGluIleHisValHisLeuSerThrAlaPheAla 260  
Db 730 GTGGAGGCCACCGTTCCTCCGCGAGATCCAGTGTTCACCTCAGCACCGCTTTGCC 789  
QY 261 ArgValAspGluAlaLeuGlyArgProGlyGlyLeuAlaValLeuAlaPheLeuGlu 280  
Db 790 AGAGTTGACGAGGCTTTGGGGCGCCGGAGGCTGGCCGTGTGGCCGCTTTCTGGAG 849  
QY 281 GluGlyProGluGluAsnSerAlaTyrGluGlnLeuLeuSerArgLeuGluIleAla 300  
Db 850 GAGGCGCCGAAGAAACAGTGCCTATGACAGTTGCTGTCTGCTTGGAAAGATCGCT 909  
QY 301 GluGluGlySerGluThrGlnValProGlyLeuAspIleSerAlaLeuLeuProSerAsp 320  
Db 910 GAGGAGGCTCAGAGACTCAGTCCAGGACTGGACATATCTGCACCTCCTGCCCTCTGAC 969  
QY 321 PheSerArgTyrPheGlnTyrGluGlySerLeuThrThrProCysAlaGlnGlyVal 340  
Db 970 TTCAGCGCTACTTCCAAATATGAGGGTCTCTGACTACACCGCCCTGTGCGCCAGGGTGC 1029  
QY 341 IleTTPThrValPheAsnGlnThrValMetLeuSerAlaLysGlnLeuHisThrLeuSer 360  
Db 1030 ATCTGGACTGTGTTTAAACGACAGTATGCTGAGTCTAAGCAGCTCCACACCTCTCT 1089

QY 361 AspThrLeuTTPGlyProGlyAspSerArgLeuGlnLeuAsnPheArgAlaThrGlnPro 380  
Db 1090 GACACCTGTGGGGACCTGGTGACTCTGGCTACAGCTGAACCTTCGAGCGACGCGCCT 1149  
QY 381 LeuAsnGlyArgValIleGluAlaSerPheProAlaGlyValAspSerProArgAla 400  
Db 1150 TTGAATGGGCGAGTGATGAGGCTCTCTCCCTGTGCTGGAGTGACAGCAGTCTCGGGCT 1209  
QY 401 AlaGluProValGlnLeuAsnSerCysLeuAlaAlaGlyAspIleLeuAlaLeuValPhe 420  
Db 1210 GCTGAGCCAGTCCAGCTGAATTCCTGCTGGCTGCTGGTGACATCTAGGCCCTGGTTTT 1269  
QY 421 GlyLeuLeuPheAlaValThrSerValAlaPheLeuValGlnMetArgArgGlnHisArg 440  
Db 1270 GGCCTCCTTTTGTCTGCACCGCTCGGGTCTCTTGTGCAGATGAGAGCGACACAGA 1329  
QY 441 ArgGlyThrIysGlyValSerTyrArgProAlaGluValAlaGluThrGlyAla 459  
Db 1330 AGGGGAACCAAGGGGTGTGAGCTACCGCCCGACGAGGTAGCCGAGACTGGAGCC 1386

RESULT 2  
US-09-772-719-1  
; Sequence 1, Application US/09772719  
; Patent No. US20020137910A1  
; GENERAL INFORMATION:  
; APPLICANT: Zavada, Jan  
; APPLICANT: Pastorekova, Silvia  
; APPLICANT: Pastorek, Jaromir  
; TITLE OF INVENTION: MN Gene and Protein  
; NUMBER OF SEQUENCES: 86  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Leona L. Lauder  
; STREET: 369 Pine Street  
; CITY: San Francisco  
; STATE: California  
; COUNTRY: USA  
; ZIP: 94104  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent Release #1.0, Version #1.30 (EPO)  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/772,719  
; FILING DATE: 30-JAN-2001  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/485,049  
; FILING DATE: 07-JUN-1995  
; APPLICATION NUMBER: US 08/260,190  
; FILING DATE: 15-JUN-1994  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Lauder, Leona L.  
; REGISTRATION NUMBER: 30,863  
; REFERENCE/DOCKET NUMBER: D-0021.3E  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 415-981-2034  
; TELEFAX: 415-981-0332  
; INFORMATION FOR SEQ ID NO: 1:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1522 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: cDNA  
; HYPOTHEetical: NO  
; ANTI-SENSE: NO  
US-09-772-719-1

Alignment Scores: 1 31e-227 Length: 1522  
Pred. No.: 2424.00 Matches: 459  
Score:



Db 133 CCCCGAGTGCAGGAGGATTCCTCCCTTGGGAGGAGGCTCTTCTTGGGGAGAGTGAACCCACTG 192  
Qy 61 GlyGluGluAspLeuProSerGluGluAspSerProArgGluGluAspProGlyGlu 80  
Db 193 GCGGAGGAGGATCTGCCAGTGAAGAGATTCACCCAGAGAGGAGGATCCACCCGGAGAG 252  
Qy 81 GluAspLeuProGlyGluGluAspLeuProGlyGluGluAspLeuProGluValLysPro 100  
Db 253 GAGGATCTACCTGAGAGGAGGATCTACCTGGAGAGGAGGATCTACCTGAAGTTAAGCCT 312  
Qy 101 LysSerGluGluGlySerLeuLysLeuGluAspLeuProThrValGluAlaProGly 120  
Db 313 AAATCAGAAGAGGAGGCTCCCTGAAGTTAGAGATCTACTACTGTGAGGCTCTCTGGA 372  
Qy 121 AspProGlnGluProGlnAsnAsnAlaHisArgAspLysGluGlyAspAspGlnSerHis 140  
Db 373 GATCCTCAAGACCCAGAGATATGCCCAAGGACCAAGAGGGGATGACACAGATCAT 432  
Qy 141 TrpArgTyrGlyGlyAspProTrpProArgValSerProAlaCysAlaGlyArgPhe 160  
Db 433 TGGCGCTATGGAGCGCACCCGCCCTGGCCCGGGTGTCCCGAGCCTGCGCGGGCGGCTTC 492  
Qy 161 GluSerProValAspIleArgProGlnLeuAlaPheCysProAlaLeuArgProLeu 180  
Db 493 CAGTCCCGGTGGATATCCCGCCCCAGCTCGCCGCTTCTGCCCGGCCCTGCGGCCCTG 552  
Qy 181 GluLeuLeuGlyPheGlnLeuProLeuProGluLeuArgLeuArgAsnAsnGlyHis 200  
Db 553 GAACCTCTGGGCTTCAGCTCCCGCCGCTCCAGAGCTGGCGCTCGCAACATGGCCAC 612  
Qy 201 SerValGlnLeuThrLeuProProGlyLeuGluMetAlaLeuGlyProGlyArgGluTyr 220  
Db 613 AGTGTCAACTGACCTTCCTCTGGGTAGAGATGCTCTGGTCCCGGGCGGAGTAC 672  
Qy 221 ArgAlaLeuGlnLeuHisLeuHisTrpGlyAlaAlaGlyArgProGlySerGluHisThr 240  
Db 673 CGGGCTCTGAGCTGCATCTGCATCGGGGGCTGCAGGTCTGTCGGGCTCGAGACACAT 732  
Qy 241 ValGluGlyHisArgPheProAlaGluIleHisValValHisLeuSerThrAlaPheAla 260  
Db 733 GTGGAGGCCACCGTTTCCCTCCGAGATCCACGTGGTTCCACCTCAGCACCGGCTTTGCC 792  
Qy 261 ArgValAspGluAlaLeuGlyArgProGlyGlyLeuAlaValLeuAlaAlaPheLeuGlu 280  
Db 793 AGAGTTACAGGAGGCTTTGGGGCGCCCGGAGGCTGGCCGCTGTTGGCGGCTTTCTGGAG 852  
Qy 281 GluGlyProGluGluAsnSerAlaTyrGluGlnLeuLeuSerArgLeuGluIleAla 300  
Db 853 GAGGGCCCGAAGAAACAGTGCCCTATGAGCAGTTGCTGTCTGCTTGGGAAGAAATCGCT 912  
Qy 301 GluGluGlySerGluThrGlnValProGlyLeuAspIleSerAlaLeuLeuProSerAsp 320  
Db 913 GAGGAGGCTCAGAGACTCAGTCCAGGCTCGAGATCTGCATCTCTGCCCTCTGAC 972  
Qy 321 PheSerArgTyrPheGlnTyrGluGlySerLeuThrThrProProCysAlaGlnGlyVal 340  
Db 973 TTCAGCGCTACTTCCAATATGAGGGTCTCTGACTACACCGGCTCTGTGCCCAGGGGTGC 1032  
Qy 341 IleTrpThrValPheAsnGlnThrValMetLeuSerAlaLysGlnLeuHisThrLeuSer 360  
Db 1033 ATCTGGAGCTGTGTTAACCAAGACAGATGATGAGTGCTAAGCAGCTCTCCACCCCTCTCT 1092  
Qy 361 AspThrLeuTrpGlyProGlyAspSerArgLeuGlnLeuAsnPheArgAlaThrGlnPro 380  
Db 1093 GACACCTGTGGGACCTGTGACTCTCGCTACAGCTGAATCTCCGAGCGACGAGCCT 1152  
Qy 381 LeuAsnGlyArgValIleGluAlaSerPheProAlaGlyValAspSerSerProArgAla 400  
Db 1153 TTGAATGGGGAGTGATTTGAGGCTCTCTTCCTGCTGGAGTGACAGCAGCTCTCGGGCT 1212  
Qy 401 AlaGluProValGlnLeuAsnSerCysLeuAlaAlaGlyAspIleLeuAlaLeuValPhe 420  
Db 1213 GCTGAGCCAGTCCAGCTGAATTTCTTGCTGGCTGCTGGTGACATCTCTAGCCCTTGTTTTT 1272

## RESULT 4

US-10-723-795-1  
; Sequence 1, Application US/10723795  
; Publication No. US20040146955A1  
; GENERAL INFORMATION:  
; APPLICANT: Supuran, Claudiu  
; APPLICANT: Scozzafava, Andrea  
; APPLICANT: Pastorekova, Silvia  
; APPLICANT: Pastorek, Jaromir  
; TITLE OF INVENTION: CA IX-SPECIFIC INHIBITORS  
; FILE REFERENCE: MST-2393 US  
; CURRENT APPLICATION NUMBER: US/10/723,795  
; CURRENT FILING DATE: 2003-11-26  
; PRIOR APPLICATION NUMBER: 60/429,089  
; PRIOR FILING DATE: 2002-11-26  
; PRIOR APPLICATION NUMBER: 60/489,473  
; PRIOR FILING DATE: 2003-07-22  
; PRIOR APPLICATION NUMBER: 60/515,140  
; PRIOR FILING DATE: 2003-10-28  
; NUMBER OF SEQ ID NOS: 9  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 1  
; LENGTH: 1522  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (13)..(1389)  
; FEATURE:  
; NAME/KEY: mat peptide  
; LOCATION: (124)..(1389)  
US-10-723-795-1

## Alignment Scores:

Pred. No.:	1,31e-227	Length:	1522
Score:	2424.00	Matches:	459
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	100.0%	Indels:	0
DB:	7	Gaps:	0

US-09-967-237B-2 (1-459) x US-10-723-795-1 (1-1522)

Qy 1 MetAlaProLeuCysProSerProTrpLeuProLeuLeuIleProAlaProAlaProGly 20  
Db 13 ATGGCTCCCTCTGTCGCCCCAGCCCTCGCTCCCTCTGTTGATCCGGCCCTCTGCTCCAGGC 72  
Qy 21 LeuThrValGlnLeuLeuLeuSerLeuLeuLeuMetProValHisProGlnArgLeu 40  
Db 73 CTCACCTGTGCAACTGCTGCTGTCTCCTCTGATGCTCTCATCCCCAGAGGTTG 132  
Qy 41 ProArgMetGlnGluAspSerProLeuGlyGlySerSerGlyGluAspAspProLeu 60  
Db 133 CCCCGATGCAGGAGGATTTCCCTTTGGGAGGAGGCTCTTCTGGGAAGATGACCCACTG 192  
Qy 61 GlyGluGluAspLeuProSerGluGluAspSerProArgGluGluAspProProGlyGlu 80  
Db 193 GGCAGAGGAGATCTGCCAGTGAAGAGATTACCCAGAGAGAGGATCCACCCGGAGAG 252  
Qy 81 GluAspLeuProGlyGluGluAspLeuProGlyGluGluAspLeuProGluValLysPro 100  
Db 253 GAGGATCTACCTGAGAGGAGGATCTACCTGGAGAGGAGGATCTACCTGAAGTTAAGCCT 312  
Qy 101 LysSerGluGluGlySerLeuLysLeuGluAspLeuProThrValGluAlaProGly 120

```
Db 313 AAATCAGAGAGAGGGGCTCCCTGAAGCTTACAGGATCTACTGTTGAGGCTCTCTGGA 372
Qy 121 AspProGlnGluProGlnAsnAsnAlaHisArgAspLysGluGlyAspAspGlnSerHis 140
Db 373 GATCCTCAAGAACCCCAAGATAATGCCACAGGACAAAGAGGGATGACAGAGTCAT 432
Qy 141 TrpArgTyrGlyGlyAspProTrpProArgValSerProAlaCysAlaGlyArgPhe 160
Db 433 TGGCGCTATGAGGCGACCCGCTTGGCCCGGGTGTCCAGGCTCGCGGGCGCGCTTC 432
Qy 161 GlnSerProValAspLysArgProGlnLeuAlaAlaPheCysProAlaLeuArgProLeu 180
Db 493 CAGTCCCGGTGATATCCGCCCCAGCTCGCGGCTTCTGCGGGCCCTGCGCCCCCTG 552
Qy 181 GluLeuLeuGlyPheGlnLeuProProLeuProGlnLeuArgLeuArgAsnGlyHis 200
Db 553 GAATCTCTGGGCTTCCAGCTCCCGCGCTCCCAAGACTCGGCGCTCGCGCAACAATGGCCAC 612
Qy 201 SerValGlnLeuThrLeuProProGlyLeuGluMetAlaLeuGlyProGlyArgGluTyr 220
Db 613 AGTGTGCACTGACCTTGCCTCTGGGCTAGAGATGGCTCTGGGTCCCGGGCGGGAGTAC 672
Qy 221 ArgAlaLeuGlnLeuHisLeuHisTrpGlyAlaAlaGlyArgProGlySerGluHisThr 240
Db 673 CGGCTCTGCAGCTGCATCTGCATCGGGGGCTGCAGGTCTCGGGCTCGAGCACACT 732
Qy 241 ValGluGlyHisArgPheProAlaGluLeuHisValHisLeuSerThrAlaPheAla 260
Db 733 GTGGAAGCCACCGCTTCCCTCGCGGAGATCCAGCTGGTGTTCACCTCAGCACCGCTTTGCC 792
Qy 261 ArgValAspGluAlaLeuGlyArgProGlyGlyLeuAlaValLeuAlaAlaPheLeuGlu 280
Db 793 AGAGTTGACAGGCGCTTGGGGCGCGCGGAGGCTCGCGGCTTGTGGCGCTTCTGGAG 852
Qy 281 GluGlyProGluGlnAsnSerAlaTyrGluGlnLeuLeuSerArgLeuGluGluLeuAla 300
Db 853 GAGGCGCGGAGAAACAGTGCCTATGAGCAGTTGCTGTCTGCTTGGGAAGAATCGCT 912
Qy 301 GluGluGlySerGluThrGlnValProGlyLeuAspLysSerAlaLeuLeuProSerAsp 320
Db 913 GAGGAAGGCTCAGAGACTCAGGTCCAGAGCTGAGCATATCTGCACCTCTCCGCTCTGAC 972
Qy 321 PheSerArgTyrPheGlnTyrGluGlySerLeuThrThrProCysAlaGlnGlyVal 340
Db 973 TTCAGCGCTACTTCCCAATATAGGGGTCTCTGACTACCGCCCTGTGCCAGGGGTGTC 1032
Qy 341 IleTrpThrValPheAsnGlnThrValMetLeuSerAlaLysGlnLeuHisThrLeuSer 360
Db 1033 ATCTGGACTGTGTTTAAACAGACAGTGTATGCTGAGTCTAAGCAGCTCCACACCTCTCT 1092
Qy 361 AspThrLeuTrpGlyProGlyAspSerArgLeuGlnLeuAsnPheArgAlaThrGlnPro 380
Db 1093 GACACCTGTGGGACCTGTGTACTCTCGGCTACAGCTGAACCTTCCGAGCGAGCGACCT 1152
Qy 381 LeuAsnGlyArgValIleGluAlaSerPheProAlaGlyValAspSerSerProArgAla 400
Db 1153 TTGAATGGGCGAGTGTATGAGGCTCTCTTCCCTCTGGAGTGACACAGTCTCTCGGGCT 1212
Qy 401 AlaGluProValGlnLeuAsnSerCysLeuAlaAlaGlyAspIleLeuAlaLeuValPhe 420
Db 1213 GCTGAGCAGTCCAGCTGMAATCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1272
Qy 421 GlyLeuLeuPheAlaValThrSerValAlaPheLeuValGlnMetArgGlnHisArg 440
Db 1273 GGCCTCTTTTGTGTGTCACAGCGTCCGCTTCTTGTGAGATGAGAGGAGCAGACAGA 1332
Qy 441 ArgGlyThrLysGlyValSerTyrArgProAlaGluValAlaGluThrGlyVala 459
Db 1333 AGGGAAACCAAGGGGTGTGAGTACCGCCAGCAGAGTAGCCGAGACTGGAGCC 1389
RESULT 5
US-10-795-933-5
```

```
; Sequence 5, Application US/10795933
; Publication No. US20040259126A1
; GENERAL INFORMATION:
; APPLICANT: Zavada, Jan
; APPLICANT: Pastorekova, Silvia
; APPLICANT: Pastorek, Jaromir
; TITLE OF INVENTION: MN Gene and Protein
; FILE REFERENCE: D-0021-2
; CURRENT APPLICATION NUMBER: US/10/795,933
; PRIOR FILING DATE: 2004-03-08
; PRIOR APPLICATION NUMBER: US/08/260,190
; PRIOR FILING DATE: 1994-06-15
; PRIOR APPLICATION NUMBER: 08/177,093
; PRIOR FILING DATE: 1993-12-30
; PRIOR APPLICATION NUMBER: 07/964,589
; PRIOR FILING DATE: 1992-10-21
; PRIOR APPLICATION NUMBER: PV-709-92
; PRIOR FILING DATE: 1992-03-11
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 5
; LENGTH: 1522
; TYPE: DNA
; ORGANISM: HUMAN
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (13)..(1389)
; FEATURE:
; NAME/KEY: mat_peptide
; LOCATION: (124)..(1389)
US-10-795-933-5
Alignment Scores:
Pred. No.: 1,31e-227 Length: 1522
Score: 2424.00 Matches: 459
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 8 Gaps: 0
US-09-967-237B-2 (1-459) x US-10-795-933-5 (1-1522)
```

```
Qy 1 MetAlaProLeuCysProSerProTrpLeuProLeuLeuLeuProAlaProGly 20
Db 13 ATGGCTCCCTGTGCTGCCAGCCCTGGCTCCCTCTGTGTGATCCGCGCCCTGTCTCCAGGC 72
Qy 21 LeuThrValGlnLeuLeuSerLeuLeuLeuMetProValHisProGlnArgLeu 40
Db 73 CTCACTGTGCACTGTGCTGTCTCACTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCT 132
Qy 41 ProArgMetGlnGluAspSerProLeuGlyGlySerSerGlyGlyGluAspProLeu 60
Db 133 CCCCGATGAGGAGGATTTCCCTTGGGAGGAGGCTCTTCTGGGAAGATGACCCACTG 192
Qy 61 GlyGluGluAspLeuProSerGluGluAspSerProArgGluGluAspProProGlyGlu 80
Db 193 GCGCAGGAGGATCTGCCAGTGAAGAGGATTCAACCCAGAGAGGAGATCCACCCGAGAG 252
Qy 81 GluAspLeuProGlyGluGluAspLeuProGlyGluGluAspLeuProGluValLysPro 100
Db 253 GAGGATCTACCTGAGAGGAGGATCTACCTGAGAGGAGGATCTACCTGAAAGTTAAGCCT 312
Qy 101 LysSerGluGluGluGlySerLeuLysLeuGluAspLeuProThrValGluAlaProGly 120
Db 313 AAATCAGAAAGAGGGCTCCCTGAAAGTTAGAGGATCTACTACTGTGTAGGCTCTCTGGA 372
Qy 121 AspProGlnGluProGlnAsnAsnAlaHisArgAspLysGlyAspAspGlnSerHis 140
Db 373 GATCCTCAGAACCCCAAGATAATGCCACAGGACAAAGAGGGATGACAGAGTCAT 432
Qy 141 TrpArgTyrGlyGlyAspProTrpProArgValSerProAlaCysAlaGlyArgPhe 160
Db 433 TGGCGCTATGAGGCGACCCGCTTGGCCCGGGTGTCCCGAGCTTCCCGGGCGCGCTTC 492
```

```
QY 161 GlnSerProValAspIleArgProGlnLeuAlaAAlaPheCysProAlaLeuArgProLeu 180
DB 493 CAGTCCCGGTGGATATCCGCCCCAGCTCGCCGCTTCTGCGCGCCCTGCGCCCCCTG 552
QY 181 GluLeuLeuGlyPheGlnLeuProProLeuProGlnLeuArgLeuArgAsnGlyHis 200
DB 553 GAACCTCTGGGCTTCCAGCTCCCGCGCTCCAGAACTGGCGCTGCGCAACAATGGCCAC 612
QY 201 SerValGlnLeuThrLeuProProGlyLeuGluMetAlaLeuGlyProGlyArgGluTyr 220
DB 613 AGTGTGCAACTGACCCCTGCCTCCTGGGCTAGAGATGGCTCTGGGTCCCGGCGGGAGTAC 672
QY 221 ArgAlaLeuGlnLeuHisLeuHisThrGlyAlaAlaGlyArgProGlySerGluHisThr 240
DB 673 CGGGCTCTGAGCTGCATCTGCATCGGGGGCTGCAGGTCTGTCGGGCTCGAGGACACT 732
QY 241 ValGluGlyHisArgPheProAlaGluIleHisValValHisLeuSerThrAlaPheAla 260
DB 733 GTGAAGGCCACCGTTTCCCTCGCGAGATCCAGTGTTCACCTCAGCACCGCTTTGGCC 792
QY 261 ArgValAspGluAlaLeuGlyArgProGlyGlyLeuAlaValLeuAlaAlaPheLeuGlu 280
DB 793 AGAGTTGACGAGCCTTGGGGCGCCCGGGAGGCTGGCCGTGTGGCGGCTTCTGGAG 852
QY 281 GluGlyProGluGlnAsnSerAlaTyrGluGlnLeuSerArgLeuGluGluIleAla 300
DB 853 GAGGGCCCGAAGAAAACAGTGGCTATGAGCAGTTGCTGTCTGCTTGGAAAGAAATCGCT 912
QY 301 GluGluGlySerGluThrGlnValProGlyLeuAspIleSerAlaLeuLeuProSerAsp 320
DB 913 GAGAAGGCTCAGAGACTCAGGTCCAGGAGTGCACATATCTGCACCTCCGCCCTCTGAC 972
QY 321 PheSerArgTyrPheGlnTyrGluGlySerLeuThrThrProCysAlaGlnGlyVal 340
DB 973 TTCAGCGCTACTTCCAAATAGGGGTCTCTGACTACACCGCCCTGTGCCAGGGTGTG 1032
QY 341 IleThrThrValPheAsnGlnThrValMetLeuSerAlaGlyGlnLeuHisThrLeuSer 360
DB 1033 ATCTGGACTGTGTTAAACAGACAGTGTGCTGAGTGTAAAGCAGCTCCACACCCCTCT 1092
QY 361 AspThrLeuTrpGlyProGlyAspSerArgLeuGlnLeuAsnPheArgAlaThrGlnPro 380
DB 1093 GACACCTGTGGGACCTGTGACTCTCGCTACAGCTGAACCTTCGAGGCGACGACCT 1152
QY 381 LeuAsnGlyArgValIleGluAlaSerPheProAlaGlyValAspSerSerProArgAla 400
DB 1153 TTGAATGGCGAGTGTAGGCTCTCTCCCTGCTGGAGTGACAGCAGTCTCTCGGCT 1212
QY 401 AlaGluProValGlnLeuAsnSerCysLeuAlaAlaGlyAspIleLeuAlaLeuValPhe 420
DB 1213 GCTGAGCCAGTCCAGCTGAATCTCGCTGGCTGTGCTGACATCTAGCCCTGGTTTTT 1272
QY 421 GlyLeuLeuPheAlaValThrSerValAlaPheLeuValGlnMetArgArgGlnHisArg 440
DB 1273 GGCTCTCTTTTGTGTCACAGCTGCGCTTCTTTGTGAGATGAGAGGCGACAGAGA 1332
QY 441 ArgGlyThrLysGlyGlyValSerTyrArgProAlaGluValAlaGluThrGlyVala 459
DB 1333 AGGGGAACCAAGGGGTGTGAGCTACCGCCAGCAGAGTAGCCGAGACTGGAGCC 1389
```

## RESULT 6

US-10-888-694-1

; Sequence 1, Application US/10888694

; Publication No. US2005003425A1

; GENERAL INFORMATION:

; APPLICANT: Zavada, Jan

; Pastorekova, Silvia

; Pastorek, Jaromir

; TITLE OF INVENTION: MN Gene and Protein

; NUMBER OF SEQUENCES: 86

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Leona L. Lauder

```
; STREET: 465 California Street, Suite 450
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30 (BPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/888,694
; FILING DATE: 08-Jul-2004
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/09/772,719
; FILING DATE: 30-Jan-2001
; APPLICATION NUMBER: US 08/485,049
; FILING DATE: 07-JUN-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Lauder, Leona L.
; REGISTRATION NUMBER: 30,863
; REFERENCE/DOCKET NUMBER: D-0021.3A-2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-981-2034
; TELEFAX: 415-981-0332
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1522 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-10-888-694-1
```

```
Alignment Scores:
Pred. No.: 1,31e-227 Length: 1522
Score: 2424.00 Matches: 459
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 8 Gaps: 0
```

US-09-967-237B-2 (1-459) x US-10-888-694-1 (1-1522)

```
QY 1 MetAlaProLeuCysProSerProTrrProLeuLeuIleProAlaProAlaProGly 20
DB 13 ATGGCTCCCTGTGCCCGCCAGCCCTGGCTCCCTCTGTATCCCGGCCCTGTCTCAGGC 72
QY 21 LeuThrValGlnLeuLeuSerLeuLeuLeuMetProValHisProGlnArgLeu 40
DB 73 CTCACGTGCAACTGCTGTGTGTCACGTCTCTGTATGCTCTCCATCCCGAGGTTG 132
QY 41 ProArgMetGlnGluAspSerProLeuGlyGlySerSerGlyGluAspProLeu 60
DB 133 CCCCAGATCAGAGGATTTCCCTTTGGGAGGAGGCTCTTCTGGGGAAGATGACCCACTG 192
QY 61 GlyGluGluAspLeuProSerGluGluAspSerProArgGluGluAspProGlyGlu 80
DB 193 GCGGAGGAGGATCTGCCAGTAGAGGATTCACCCAGAGAGAGATCCACCCGAGAG 252
QY 81 GluAspLeuProGlyGluGluAspLeuProGlyGluGluAspLeuProGlyValLysPro 100
DB 253 GAGGATCTACCTGGAGAGGAGGATCTACCTGGAGAGGAGATCTACTGAAGTTAAGCCT 312
QY 101 LysSerGluGluGluGlySerLeuLysLeuGluAspLeuProThrValGluAlaProGly 120
DB 313 AAATCAGAAGAAGAGGCTCCCTGGAAGTTAGAGGATCTACTACTCTTTGAGGCTCTCGGA 372
QY 121 AspProGlnGluProGlnAsnAlaHisArgAspLysGluGlyAspAspGlnSerHis 140
```

```
Db 373 GATCCTCAAGAACCCAGATAATATGCCCACAGGACAAAGAGGGGATGACAGAGTCAT 432
Qy 141 TTPArgTyrGlyGlyAspProTTPProArgValSerProAlaCyAsAlaGlyArgPhe 160
Db 433 TGGCGCTATGAGGCGACCCCGCTGCGCGCGGTGTCCCGAGCGCTGCGGGCGCGCTTC 492
Qy 161 GlnSerProValAspIleArgProGlnLeuAlaAlaPheCyAsProAlaLeuArgProLeu 180
Db 493 CAGTCCCGGTGGATATCCGCCCCCAGCTGCGCGCTTCTGCGCGCGCTGCGCCCCCTG 552
Qy 181 GluLeuLeuGlyPheGlnLeuProProLeuProGlnLeuArgLeuArgAsnAnglyHis 200
Db 553 GAACCTCTGGGCTTCCAGCTCCCGCGCTCCAGAACTGCGCGCTCGCAACAATGGCCAC 612
Qy 201 SerValGlnLeuThrLeuProProGlyLeuGluMetAlaLeuGlyProGlyArgGluTyr 220
Db 613 AGTGTGAACCTGACCCCTGCTTCCGCGCTAGAGATGGCTCTGGGTCCCGGGCGGGAGTAC 672
Qy 221 ArgAlaLeuGlnLeuHisLeuHisTTPGlyAlaAlaGlyArgProGlySerGluHisThr 240
Db 673 CGGCTCTGAGCTGCAATCTGCACTGGGGGGCTGCAGGTGCTCGGGCTCGAGCACACT 732
Qy 241 ValGluGlyHisArgPheProAlaGluIleHisValValHisLeuSerThrAlaPheAla 260
Db 733 GTGGAAGCCACCGTTCCTCCGCGGAGATCCACGTGTTCACCTCAGCACCGCCTTTGCC 792
Qy 261 ArgValAspGluAlaLeuGlyArgProGlyGlyLeuAlaValLeuAlaAlaPheLeuGlu 280
Db 793 AGAGTTGACGAGGCTTTGGGGCGCGCGGAGGCGCTGGCGGTGGTGGCGCGCTTTCTGGAG 852
Qy 281 GluGlyProGluGlnAsnSerAlaTyrGluGlnLeuLeuSerArgLeuGluGluIleAla 300
Db 853 GAGGCGCGGAAGAAACAGTGCCTATGAGCAGTGTCTGCTCTGCTTGGAGAAATCGCT 912
Qy 301 GluGlySerGluThrGlnValProGlyLeuAspIleSerAlaLeuLeuProSerAsp 320
Db 913 GAGGAAGCTCAGAGACTCAGTCCAGGAGTGCACATATCTGCACCTCGCTCGCTCTGAC 972
Qy 321 PheSerArgTyrPheGlnTyrGlnGlySerLeuThrThrProProCyAsAlaGlnGlyVal 340
Db 973 TTCAGCGCTACTTCCAAATATAGGGGTCTCTGACTACACCGCGCTGTGCGCCAGGGTGC 1032
Qy 341 IleTTPThrValPheAsnGlnThrValMetLeuSerAlaLysGlnLeuHisThrLeuSer 360
Db 1033 ATCTGGAGTGTGTTTACACAGACAGTGAATGTGAGTGTGAAGCAGCTCCACACCTCTCT 1092
Qy 361 AspThrLeuTTPGlyProGlyAspSerArgLeuGlnLeuAsnPheArgAlaThrGlnPro 380
Db 1093 GACACCTGTGGGACCTGTGACTCTCGGTACAGCTGAACCTCCGAGCGACGACGCT 1152
Qy 381 LeuAsnGlyArgValIleGluAlaSerPheProAlaGlyValAspSerProArgAla 400
Db 1153 TTGAATGGCGAGTGAATGAGCGCTCTTCCCTGCTGGAGTGGACAGCAGTCTCTCGGGCT 1212
Qy 401 AlaGluProValGlnLeuAsnSerCysLeuAlaAlaGlyAspIleLeuAlaLeuValPhe 420
Db 1213 GCTGAGCAGGCCAGCTGAATTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1272
Qy 421 GlyLeuLeuPheAlaValThrSerValAlaPheLeuValGlnMetArgArgGlnHisArg 440
Db 1273 GGCTCTCTTTTGTGTGTACACAGGCTGCGGTCTCTTGTGCAGATGAGAGCGACACAGA 1332
Qy 441 ArgGlyThrIysGlyValSerTyrArgProAlaGluValAlaGluThrGlyAla 459
Db 1333 AGGGGAACCAAGGGGTGTGAGCTACCGCCAGCAGAGGTAGCCGAGACTGGAGCC 1389
```

## RESULT 7

US-10-921-590-1

; Sequence 1, Application US/10921590

; Publication No. US20050031623A1

; GENERAL INFORMATION:

; APPLICANT: Pastorek, Jaromir

```
; APPLICANT: Zavada, Jan
; APPLICANT: Ortova Gut, Marta
; APPLICANT: Zatovicova, Miriam
; APPLICANT: Pastorekova, Silvia
; APPLICANT: Zavadova, Zuzanna
; TITLE OF INVENTION: SOLUBLE FORM OF CARBONIC ANHYDRASE IX (8-CA IX), ASSAYS TO DETECT
; TITLE OF INVENTION: 8-CA IX, CA IX'S COEXPRESSION WITH HER-2/neu/c-erbB-2 AND CA
; FILE REFERENCE: IX-SPECIFIC MONOCLONAL ANTIBODIES TO NON-IMMUNODOMINANT EPITOPES
; CURRENT APPLICATION NUMBER: US/10/921,590
; CURRENT FILING DATE: 2004-08-19
; PRIOR APPLICATION NUMBER: 60/358,824
; PRIOR FILING DATE: 2002-02-21
; PRIOR APPLICATION NUMBER: 60/383,068
; PRIOR FILING DATE: 2002-05-23
; PRIOR APPLICATION NUMBER: 60/431,499
; PRIOR FILING DATE: 2002-12-05
; PRIOR APPLICATION NUMBER: PCT/US03/05136
; PRIOR FILING DATE: 2003-02-21
; PRIOR APPLICATION NUMBER: PCT/US03/05137
; PRIOR FILING DATE: 2003-02-21
; NUMBER OF SEQ ID NOS: 118
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 1
; LENGTH: 1522
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (13)..(1389)
; FEATURE:
; NAME/KEY: mat peptide
; LOCATION: (124)..(1389)
; US-10-921-590-1
```

## Alignment Scores:

Pred. No.:	1,31e-227	Length:	1522
Score:	2424.00	Matches:	459
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	100.0%	Indels:	0
DB:	8	Gaps:	0

US-09-967-237B-2 (1-459) x US-10-921-590-1 (1-1522)

Qy	1	MetAlaProLeuCyAsProSerProTTPLeuProLeuLeuLeuProAlaProAlaProGly	20
Db	13	ATGGCTCCCTGTGCCCCCAGCCCTGGCTCCCTCTGTGTGATCCGCGCCCTGTCTCCAGGC	72
Qy	21	LeuThrValGlnLeuLeuSerLeuLeuLeuMetProValHisProGlnArgLeu	40
Db	73	CTACTGTGCACTGTGCTGTCTACTGCTGCTTCTGATGCTGTCCATCCCAGAGGTTG	132
Qy	41	ProArgMetGlnGluAspSerProLeuGlyGlySerSerGlyGluAspProLeu	60
Db	133	CCCGGATGACGAGGATTCCTCCCTTGGAGGAGGCTCTTCTGGGGAAGATGACCCACTG	192
Qy	61	GlyGluGluAspLeuProSerGluGluAspSerProArgGluGluAspProGlyGlu	80
Db	193	GGCGAGGAGGATTCGCCAGTGAAGAGGATTCACCCAGAGAGGAGATCCACCCGAGAG	252
Qy	81	GluAspLeuProGlyGluGluAspLeuProGlyGluGluAspLeuProGluValIysPro	100
Db	253	GAGATCTACTGGAGGAGGAGGATCTACTCTGGAGAGGAGGATCTACTGAAGTTAAGCT	312
Qy	101	LysSerGluGluGlySerLeuLeuGluAspLeuProThrValGluAlaProGly	120
Db	313	AAATCAGAAGAAGAGGCTCCCTGAAGTTAGAGGATCTACTACTGTTGAGGCTCTCGA	372
Qy	121	AspProGlnGluProGlnAsnAsnAlaHisArgAspIysGluGlyAspAspGlnSerHis	140
Db	373	GATCCTCAAGAACCCAGAAATAATGCCACAGGACAAAGAGGGGATGACACAGATCAT	432



```
Db 523 CAGTCCCGGTGGATATCCGCCCGCCAGCTCGCGCCCTTCTGCGCGCCCTGCGCCCGCTG 582
Qy 181 GluLeuLeuGlyPheGlnLeuProLeuProLeuProGluLeuArgLeuArgLeuHis 200
Db 583 GAACTCTCTGGGCTTCCAGCTCCCGCCGCTCCCAAGACTGGCGCTCGCAACAATGGCCAC 642
Qy 201 SerValGlnLeuThrLeuProGluLeuMetAlaLeuGlyProGlyArgGluTyr 220
Db 643 AGTGTGCACTGACCTTCCCTTCTGGGCTAGAGATGGCTTGGGTCCCGGCGGGAGTAC 702
Qy 221 ArgAlaLeuGlnLeuHisLeuHisTrpGlyAlaAlaGlyArgProGlySerGluHisThr 240
Db 703 CGGCTCTGCGAGCTGCATCTGCATCGCGGGCTGCAGGCTCGTCCGGCTCGGAGCACACT 762
Qy 241 ValGluGlyHisArgPheProAlaGluLeuHisValValHisLeuSerThrAlaPheAla 260
Db 763 GTGGAAGGCCACCGTTTCCCTCGCGAGATCCAGTGGTTTCCACTCAGCACCGGCTTTGCC 822
Qy 261 ArgValAspGluAlaLeuGlyArgProGlyGlyLeuAlaValLeuAlaAlaPheLeuGlu 280
Db 823 AGAGTTGACGAGGCTTTGGGGCGCCCGGAGGCTGCGCGCTTGGCGGCTTCTTGGAG 882
Qy 281 GluGlyProGluGluAsnSerAlaTyrGluGlnLeuLeuSerArgLeuGluGluLeuAla 300
Db 883 GAGGCGCGGAGAAACAGTGCCTATGAGCAGTTGCTGCTCGCTTGGAGAAATCGCT 942
Qy 301 GluGluGlySerGluThrGlnValProGlyLeuAspLeuSerAlaLeuLeuProSerAsp 320
Db 943 GAGGAGGCTCAGAGACTCAGGTCCAGGAGTGCAGATATCTGCACCTCTCCCTCTGAC 1002
Qy 321 PheSerArgTyrPheGlnTyrGluGlySerLeuThrProProCysAlaGlnGlyVal 340
Db 1003 TTCAGCGCGTACTTCCAAATAGAGGGGTCTCTGACTACCGCCCTGTGCCCGAGGGTGC 1062
Qy 341 IleTrpThrValPheAsnGlnThrValMetLeuSerAlaLysGlnLeuHisThrLeuSer 360
Db 1063 ATCTGGACTGTGTTAACACAGACAGATGATGCTGAGTCTTAAGCAGCTCCACACCTCTCT 1122
Qy 361 AspThrLeuTrpGlyProGlyAspSerArgLeuGlnLeuAsnPheArgAlaThrGlnPro 380
Db 1123 GACACCTGTGGGACCTGTGTGACTCTCGGCTACAGCTGAACCTTCCGAGGCGACGACCT 1182
Qy 381 LeuAsnGlyArgValIleGluAlaSerPheProAlaGlyValAspSerSerProArgAla 400
Db 1183 TTGAATGGCGAGTGATGAGGCTCTCTTCCCTCTGGAGTGGACAGCTCTCTCGGCT 1242
Qy 401 AlaGluProValGlnLeuAsnSerCysLeuAlaAlaGlyAspIleLeuAlaLeuValPhe 420
Db 1243 GCTGAGCCAGTCCAGCTGAATCTCTGCTGGCTGCTGCTGACATCTTAGCCCTGGTTTTT 1302
Qy 421 GlyLeuLeuPheAlaValThrSerValAlaPheLeuValGlnMetArgArgGlnHisArg 440
Db 1303 GGCCTCTTTTGTGTCAACAGCGTCCGCTTCTTCTGTCAGATGAGAGGAGCAGACAGA 1362
Qy 441 ArgGlyThrIysGlyValSerTyrArgProAlaGluValAlaGluThrGlyAla 459
Db 1363 AGGGGAACCAAGGGGGTGTGAGTACTCGCCAGCAGAGGTAGCCGAGACTGGAGCC 1419

RESULT 9
US-09-954-456-726
; Sequence 726, Application US/09954456
; Patent No. US20020115057A1
; GENERAL INFORMATION:
; APPLICANT: Young, Paul
; TITLE OF INVENTION: Process for Identifying Anti-Cancer Therapeutic Agents Using Canc
; TITLE OF INVENTION: Sets
; FILE REFERENCE: 689290-76
; CURRENT APPLICATION NUMBER: US/09/954,456
; PRIOR FILING DATE: 2001-09-18
; PRIOR APPLICATION NUMBER: US/60/233,617
; PRIOR FILING DATE: 2000-09-18
; PRIOR APPLICATION NUMBER: US/60/234,052
```

```
; PRIOR FILING DATE: 2000-09-20
; PRIOR APPLICATION NUMBER: US/60/234,923
; PRIOR FILING DATE: 2000-09-25
; PRIOR APPLICATION NUMBER: US/60/235,134
; PRIOR FILING DATE: 2000-09-25
; PRIOR APPLICATION NUMBER: US/60/235,637
; PRIOR FILING DATE: 2000-09-26
; PRIOR APPLICATION NUMBER: US/60/235,638
; PRIOR FILING DATE: 2000-09-26
; PRIOR APPLICATION NUMBER: US/60/235,711
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: US/60/235,720
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: US/60/235,840
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: US/60/235,863
; PRIOR FILING DATE: 2000-09-27
; NUMBER OF SEQ ID NOS: 2276
; SOFTWARE: Patent in version 3.0
; SEQ ID NO 726
; LENGTH: 1552
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-954-456-726

Alignment Scores:
Pred. No.: 1,346-227 Length: 1552
Score: 2424.00 Matches: 459
Percent Similarity: 100.0% Conservatives: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 3 Gaps: 0

US-09-967-237B-2 (1-459) x US-09-954-456-726 (1-1552)
Qy 1 MetAlaProLeuCysProSerProTrpLeuProLeuLeuProAlaProAlaProGly 20
Db 43 ATGGCTCCCTGTGCGCCAGCCCTGGCTCCCTCTGTGTGATCCCGGCCCTGTCTCCAGGC 102
Qy 21 LeuThrValGlnLeuLeuSerLeuLeuLeuMetProValHisProGlnArgLeu 40
Db 103 CTCACTGTGCACTGTCTGCTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 162
Qy 41 ProArgMetGlnGluAspSerProLeuGlyGlySerSerGlyGluAspAspProLeu 60
Db 163 CCCCGGATGAGGAGGATTTCCCTTTGGGAGGAGGCTCTTCTGGGAGAGATGACCCACTG 222
Qy 61 GlyGluGluAspLeuProSerGluGluAspSerProArgGluGluAspProProGlyGlu 80
Db 223 GCGCAGGAGGATCTGCCAGTGAAGAGGATTCACCCAGAGAGGAGGATCCACCCCGAGAG 282
Qy 81 GluAspLeuProGlyGluGluAspLeuProGlyGluGluAspLeuProGluValLeuPro 100
Db 283 GAGGATCTACCTGAGAGGAGGATCTACCTGGAGAGGAGGATCTTACCTGAGTTAAGCCT 342
Qy 101 LysSerGluGluGlySerLeuLeuLeuAspLeuProThrValGluAlaProGly 120
Db 343 AAATCAGAGAGAGGAGGCTCCCTGAGTTAGAGGATCTACTACTGTTGAGGCTCTCTGGA 402
Qy 121 AspProGlnGluProGlnAsnAlaHisArgAspLysGluGlyAspAspGlnSerHis 140
Db 403 GATCCTCAAGAACCCAGAGATAATCCACAGGACAAAGAGGAGGATGACACAGATCAT 462
Qy 141 TrpArgTyrGlyGlyAspProProTrpProArgValSerProAlaCysAlaGlyArgPhe 160
Db 463 TGGCGCTATGAGGAGGAGGAGGCTCCCTGAGTTAGAGGATCTACTACTGTTGAGGCTCTCTG 522
Qy 161 GlnSerProValAspIleArgProGlnLeuAlaAlaPheCysProAlaLeuArgProLeu 180
Db 523 CAGTCCCGGTGGATATCCCGCCCTTCTGCGGCTTCTGCGGCGCTTCTGCGGCGCTTCTG 582
Qy 181 GluLeuLeuGlyPheGlnLeuProProLeuProGluLeuArgLeuArgAsnAsnGlyHis 200
```



Db 883 GAGGCGCGAGAGAAACAGTCTGATGACAGTTCCTGCTGCTGGAGAAATCGCT 942  
Qy 301 GluGluGlySerGluThrGlnValProGlyLeuAspIleSerAlaLeuLeuProSerAsp 320  
Db 943 GAGGAAGGCTCAGAGACTCAGGTCCAGGAGTGGACATATCTGCATCTCTGCTGCTGAC 1002  
Qy 321 PheSerArgTyrPheGlnTyrGluGlySerLeuThrThrProCysAlaGlnGlyVal 340  
Db 1003 TTGAGCGCGTACTTCCAAATATGAGGGGTCTCTGACTACCGCCCTGTGCGGAGGTGTC 1062  
Qy 341 IleTrpThrValPheAsnGlnThrValMetLeuSerAlaLysGlnLeuHisThrLeuSer 360  
Db 1063 ATCTGGACTGTGTTAAACACAGACAGTGTCTGAGTGTCTAAGCAGCTCCACACCTCTCT 1122  
Qy 361 AspThrLeuTrpGlyProGlyAspSerArgLeuGlnLeuAsnPheArgAlaThrGlnPro 380  
Db 1123 GACACCCCTGTGGGACCTGTGTGACTCTCTGGCTACAGCTGAATTCGAGCGACGAGCCT 1182  
Qy 381 LeuAsnGlyArgValIleGluAlaSerPheProAlaGlyValAspSerProArgAla 400  
Db 1183 TTGAATGGCGAGTGTATGAGGCTCTCTCCCTGCTGGAGTGGACAGCAGTCTCTCGGCT 1242  
Qy 401 AlaGluProValGlnLeuAsnSerCysLeuAlaAlaGlyAspIleLeuAlaLeuValPhe 420  
Db 1243 GCTGAGCCAGTCCAGCTGAATTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1302  
Qy 421 GlyLeuLeuPheAlaValThrSerValAlaPheLeuValGlnMetArgArgGlnHisArg 440  
Db 1303 GGCTCTCTTTTGTGTGTCCACAGCGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1362  
Qy 441 ArgGlyThrLysGlyValSerTyrArgProAlaGluValAlaGluThrGlyAla 459  
Db 1363 AGGGGAACCAAGAGGGGTGTGAGTACCGCCAGCAGAGGTGCGGAGCTGGAGCC 1419

## RESULT 11

US-09-873-367C-516  
; Sequence 516, Application US/09873367C  
; Publication No. US20030165839A1  
; GENERAL INFORMATION:  
; APPLICANT: Young, Paul  
; APPLICANT: Soppet, Daniel  
; APPLICANT: Endress, Gregory  
; APPLICANT: Augustus, Meena  
; APPLICANT: Ebner, Reinhard  
; APPLICANT: Carter, Kenneth  
; TITLE OF INVENTION: Cancer Gene Determination and Therapeutic Screening Using  
; FILE REFERENCE: 689290-64  
; CURRENT APPLICATION NUMBER: US/09/873,367C  
; CURRENT FILING DATE: 2003-04-29  
; PRIOR APPLICATION NUMBER: U.S. 60/236,891  
; PRIOR FILING DATE: 2000-09-29  
; PRIOR APPLICATION NUMBER: U.S. 60/236,842  
; PRIOR FILING DATE: 2000-09-29  
; PRIOR APPLICATION NUMBER: U.S. 60/244,867  
; PRIOR FILING DATE: 2000-11-01  
; PRIOR APPLICATION NUMBER: U.S. 60/245,084  
; PRIOR FILING DATE: 2000-11-01  
; NUMBER OF SEQ ID NOS: 1067  
; SOFTWARE: Patent in version 3.0  
; SEQ ID NO 516  
; LENGTH: 1552  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-873-367C-516

## Alignment Scores:

Pred. No.:	1,34e-227	Length:	1552
Score:	2424.00	Matches:	459
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	100.0%	Indels:	0

DB: 3 Gaps: 0  
US-09-967-237B-2 (1-459) x US-09-873-367C-516 (1-1552)  
Qy 1 MetAlaProLeuCysProSerProTrpLeuProLeuLeuLeuProAlaProGly 20  
Db 43 ATGGCTCCCTGTGCGCCAGCCCTGGCTCTCTGTGATCCCGGCCCTGTCTCCAGGC 102  
Qy 21 LeuThrValGlnLeuLeuSerLeuLeuMetProValHisProGlnArgLeu 40  
Db 103 CTACATGTGCAACTGCTGCTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 162  
Qy 41 ProArgMetGlnGluAspSerProLeuGlyGlySerSerGlyGluAspProLeu 60  
Db 163 CCCCAGATGACAGGAGATTTCCCTTTGGGAGGAGGCTCTTCTGGGAGAGATGATGCCACTG 222  
Qy 61 GlyGluGluAspLeuProSerGluGluAspSerProArgGluGluAspProGlyGlu 80  
Db 223 GCGCAGAGGAGATCTGCCAGTGAAGAGGATTCACCCAGAGAGAGGATTCACCCGGAGAG 282  
Qy 81 GluAspLeuProGlyGluGluAspLeuProGlyGluGluAspLeuProGluValLysPro 100  
Db 283 GAGGATCTACCTGAGAGGAGGATCTACCTGGAGAGAGGATCTACCTGAGGTAAAGCCT 342  
Qy 101 LysSerGluGluGlySerLeuLysLeuGluAspLeuProThrValGluAlaProGly 120  
Db 343 AAATCAGAGAGAGAGGCTCCCTGGAAGTTAGAGGATCTACCTACTGTTGAGGCTCTCTGGA 402  
Qy 121 AspProGlnGluProGlnAsnAlaHisArgAspLysGlyGluAspAspGlnSerHis 140  
Db 403 GATCTCTCAAGAACCCAGAAATATGCCACAGGACAAAGAGGAGATGACAGAGGTAT 462  
Qy 141 TrpArgTyrGlyGlyAspProProTrpProArgValSerProAlaCysAlaGlyArgPhe 160  
Db 463 TGGCGCTATGAGAGCGACCCCGCTGGCCCGGGTGTCCCGACCTGCGCGGGCGGCTTC 522  
Qy 161 GlnSerProValAspIleArgProGlnLeuAlaAlaPheCysProAlaLeuArgProLeu 180  
Db 523 CAGTCCCGGTGGATATCCGCCCCAGCTCGCGCTCTGCGCGGCGCTGCGCCCTG 582  
Qy 181 GluLeuGlyPheGlnLeuProProLeuProGluLeuArgLeuArgAsnArgHis 200  
Db 583 GAACTCTCGGCTTCCAGCTCCCGCGCTCCCAAGATGCGCTCGCGCAATGGCCAC 642  
Qy 201 SerValGlnLeuThrLeuProProGlyLeuGluMetAlaLeuGlyProGlyArgGluTyr 220  
Db 643 AGTGTGCAACTGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 702  
Qy 221 ArgAlaLeuGlnLeuHisLeuHisTrpGlyAlaAlaGlyArgProGlySerGluHisThr 240  
Db 703 CGGGCTCTGCAGCTGCATCTGCATCGGGGGCTGCAGGTGCTCGGGCTCGGAGCACACT 762  
Qy 241 ValGluGlyHisArgPheProAlaGluIleHisValHisLeuSerThrAlaPheAla 260  
Db 763 GTGGAAGGCGCACCGTTTCCCTGCGGAGATCCACGTGTTCCACTCAGCACCGCTTTGGCC 822  
Qy 261 ArgValAspGluAlaLeuGlyArgProGlyGlyLeuAlaValLeuAlaAlaPheLeuGlu 280  
Db 823 AGAGTTGACGAGGCTTTGGGGCGCCCGGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 882  
Qy 281 GluGlyProGluGluAsnSerAlaTyrGluGlnLeuLeuSerArgLeuGluGluAla 300  
Db 883 GAGGGCCCGGAAGAAACAGTGTCTATGAGCAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCT 942  
Qy 301 GluGluGlySerGluThrGlnValProGlyLeuAspIleSerAlaLeuLeuProSerAsp 320  
Db 943 GAGGAAGGCTCAGAGACTCAGGTCCAGGACTGACATATCTGCATCTCTGCTGCTGCTGCTGCT 1002  
Qy 321 PheSerArgTyrPheGlnTyrGluGlySerLeuThrThrProCysAlaGlnGlyVal 340  
Db 1003 TTCAGCGCTACTTCAATATGAGGGGTCTCTGACTACACCGCCCTGTGCGGAGGTGTC 1062  
Qy 341 IleTrpThrValPheAsnGlnThrValMetLeuSerAlaLysGlnLeuHisThrLeuSer 360

```
Db 1063 ATCTGACGTGTGTAAACACAGAGTGTGCTGAGTGCTAAGCAGCTCCACACCCCTCTT 1122
Qy 361 AspThrLeuTyrGlyProGlyAspSerArgLeuGlnLeuAsnPheArgAlaThrGlnPro 380
Db 1123 GACACCTGTGGGACCTGTGTACTCTCGGTACAGCTGNACTTCGAGGCGAGCAGCCT 1182
Qy 381 LeuAsnGlyArgValIleGluAlaSerPheProAlaGlyValAspSerSerProArgAla 400
Db 1183 TTGAATGGGCGAGTGTAGGCTCTCTTCCCTGCTGGAGTGACAGCAGCTCTCGGGCT 1242
Qy 401 AlaGluProValGlnLeuAsnSerCysLeuAlaAlaGlyAspIleLeuAlaLeuValPhe 420
Db 1243 GCTGAGCCAGTCCAGCTGAATCTCTCGCTGGCTGGCTGGTGACATCTAGCCCTGGGTTTTT 1302
Qy 421 GlyLeuLeuPheAlaValThrSerValAlaPheLeuValGlnMetArgArgGlnHisArg 440
Db 1303 GGCTCTCTTTTGTCTGTACCAGCTGGCTTCTTGTGCGAGATGAGAGGCGAGCAGCA 1362
Qy 441 ArgGlyThrIysGlyGlyValSerTyrArgProAlaGluValAlaGluThrGlyAla 459
Db 1363 AGGGGAACCAAGGGGTGTGAGCTACCGCCAGCAGAGGTAGCCGAGACTGGAGCC 1419

RESULT 12
US-09-968-007A-213
; Sequence 213, Application US/09968007A
; Publication No. US20040115625A1
; GENERAL INFORMATION:
; APPLICANT: Ebner, Reinhard
; TITLE OF INVENTION: Cancer Gene Determination and Therapeutic Screening Using Signa
; FILE REFERENCE: 689290-71
; CURRENT FILING DATE: 2001-10-02
; PRIOR FILING DATE: 2001-10-02
; PRIOR APPLICATION NUMBER: US/60/237,172
; PRIOR FILING DATE: 2000-10-02
; PRIOR APPLICATION NUMBER: US/60/237,173
; PRIOR FILING DATE: 2000-10-02
; PRIOR APPLICATION NUMBER: US/60/237,278
; PRIOR FILING DATE: 2000-10-02
; PRIOR APPLICATION NUMBER: US/60/237,294
; PRIOR FILING DATE: 2000-10-02
; PRIOR APPLICATION NUMBER: US/60/237,295
; PRIOR FILING DATE: 2000-10-02
; PRIOR APPLICATION NUMBER: US/60/237,316
; PRIOR FILING DATE: 2000-10-02
; NUMBER OF SEQ ID NOS: 1001
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 213
; LENGTH: 1552
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-968-007A-213

Alignment Scores:
Pred. No.: 1,34e-227 Length: 1552
Score: 2424.00 Matches: 459
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 3 Gaps: 0

US-09-967-237B-2 (1-459) x US-09-968-007A-213 (1-1552)
Qy 1 MetAlaProLeuCysProSerProTyrLeuProLeuLeuIleProAlaProAlaProGly 20
Db 43 ATGGCTCCCTGTGGCCCGCCCGCTGGCTCCCTCTGTGTGATCCCGGCCCTGTCTCCAGGC 102
Qy 21 LeuThrValGlnLeuLeuSerLeuLeuLeuMetProValHisProGlnArgLeu 40
Db 103 CTCACTGTGCAACTGTGCTGTCTCACTGCTCTCTGATGCTGTCCATCCCGAGAGGTTG 162
Qy 41 ProArgMetGlnGluAspSerProLeuGlyGlySerSerGlyGluAspProLeu 60
```

```
Db 163 CCCCGGATGTCAGAGGATTCCTCCCTTTGGAGGAGGCTCTTCTGGGGAAGATGACCCACTG 222
Qy 61 GlyGluGluAspLeuProSerGluGluAspSerProArgGluGluAspProGlyGlu 80
Db 223 GCGGAGGAGGATCTGCCAGTGAAGGATTCACCCAGAGAGGAGGATCCACCCGAGAG 282
Qy 81 GluAspLeuProGlyGluGluAspLeuProGlyGluGluAspLeuProGluValIysPro 100
Db 283 GAGGATCTACCTGCGAGAGGAGGATCTACCTGAGAGGAGGATCTACCTGAAGTTAGCCT 342
Qy 101 LysSerGluGluGlySerLeuIysLeuGluAspLeuProThrValGluAlaProGly 120
Db 343 AAATCAGAAGAAGAGGCTCCCTCGAAGTTAGAGGATCTACCTACTGTTGAGGCTCTCTGA 402
Qy 121 AspProGlnGluProGlnAsnAlaHisArgAspLysGluGlyAspAspGlnSerHis 140
Db 403 GATCTCAGAACCCCGAGATATATGCCACACAGGACAAAGAGGGGATGACACAGATCAT 462
Qy 141 TrpArgTyrGlyGlyAspProProTyrProArgValSerProAlaCysAlaGlyArgPhe 160
Db 463 TGGCGCTATGAGGCGGACCCGCGCTGGCCCGGGTGTCCCGAGCCTGCGCGGGCGCTTC 522
Qy 161 GlnSerProValAspIleArgProGlnLeuAlaAlaPheCysProAlaLeuArgProLeu 180
Db 523 CAGTCCCGGTGGATATCCGCCCGCAGCTGCGCCCTTCTGCCCGGCCCTGCGCCCCCTG 582
Qy 181 GluLeuLeuGlyPheGlnLeuProProLeuProGluLeuArgLeuArgAsnGlnHis 200
Db 583 GAACTCTGGGCTTCCAGCTCCCGCGCTCCCGAGAACTGCGCTGCGCAACAATGSCCAG 642
Qy 201 SerValGlnLeuThrLeuProProGlyLeuGluMetAlaLeuGlyProGlyArgGluTyr 220
Db 643 AGTGTGCAACTGACCTGCTCTCGGGCTAGAGATGGCTCTCGGGTCCCGGGCGGGAGTAC 702
Qy 221 ArgAlaLeuGlnLeuHisLeuHisTrpGlyAlaAlaGlyArgProGlySerGluHisThr 240
Db 703 CGGGCTCTGCAGCTGCATCTGCACCTGGGGGGCTCGAGGTGCTCCGGGCTCGGAGCAGCT 762
Qy 241 ValGluGlyHisArgPheProAlaGluIleHisValValHisLeuSerThrAlaPheAla 260
Db 763 GTGGAGGCGCACCGTTTCCCTCCCGAGATCCACGTGTTTCACTCAGCACCGCTTTTGGC 822
Qy 261 ArgValAspGluAlaLeuGlyArgProGlyGlyLeuAlaValLeuAlaAlaPheLeuGlu 280
Db 823 AGAGTTGACAGAGGCTTTGGGGCGCCCGGAGGCTGCGCGCTGTGTGGCCCTTTCTGGAG 882
Qy 281 GluGlyProGluGluAsnSerAlaTyrGluGlnLeuLeuSerArgLeuGluGluIleAla 300
Db 883 GAGGGCCCGGAAGAAACAGTGCCTATGAGCAGTGTCTGTCTGCTTGGAAAGAAATCGCT 942
Qy 301 GluGluGlySerGluThrGlnValProGlyLeuAspIleSerAlaLeuLeuProSerAsp 320
Db 943 GAGGAAGGCTCAGAGACTCAGGTCCCGAGACTGGACATATCTGCATCTCTCGCCCTCTGAC 1002
Qy 321 PheSerArgTyrPheGlnTyrGluGlySerLeuThrThrProProCysAlaGlnGlyVal 340
Db 1003 TTCAGCCGCTACTTCCAAATATGAGGGGTCTCTGACTACACCGCCCTGTGCCAGGGTGC 1062
Qy 341 IleTrpThrValPheAsnGlnThrValMetLeuSerAlaLysGlnLeuHisThrLeuSer 360
Db 1063 ATCTGACCTGTGTTTAACACAGACAGTATGCTGAGTGTAGCAGGTCTCCACACCTCTCT 1122
Qy 361 AspThrLeuTyrGlyProGlyAspSerArgLeuGlnLeuAsnPheArgAlaThrGlnPro 380
Db 1123 GACACCTGTGGGACCTGCTGACTCTCGGCTACAGTGAACCTTCGAGCGACGAGCCT 1182
Qy 381 LeuAsnGlyArgValIleGluAlaSerPheProAlaGlyValAspSerSerProArgAla 400
Db 1183 TTGAATGGGCGAGTGTAGGCTCTCTTCCCTGCTGGAGTGGAGCAGAGTCTCTCGGGCT 1242
Qy 401 AlaGluProValGlnLeuAsnSerCysLeuAlaAlaGlyAspIleLeuAlaLeuValPhe 420
```







**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.7  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model  
Run on: February 17, 2006, 15:32:07 ; Search time 518 Seconds  
(without alignments)  
1880.959 Million cell updates/sec

Title: US-09-967-237B-2

Perfect score: 2424

Sequence: 1 MAPLCPSPWLPPLIPAPAG.....RRRTKGVSYPRAEVAETGA 459

Scoring table:

BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 7204252 seqs, 1061369211 residues

Total number of hits satisfying chosen parameters: 14408504

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlp  
-Q=/abes/ABSSWEB.spool/US0967237/runat\_15022006\_182256\_28276/app\_query.fasta\_1  
-DB=Published Applications NA New -QFMT=fastap -SUFFIX=rnpbn -MINMATCH=0.1  
-LOOPCL=0 -LOPEXT=0 -UNITS=bites -START=1 -END=1 -MATRIX=blosum62  
-TRANS=human40.cgi -LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100  
-THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0  
-MAXLEN=2000000000 -HOST=abss06p  
-USER=US0967237 @CGN 1.1 335 @runat\_15022006\_182256\_28276 -NCPU=6 -ICPU=3  
-NO MMAP -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG -DEV TIMEOUT=120  
-WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7  
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

Published Applications NA New:  
1: /cgn2\_6/ptodata/1/pubpna/US08\_NEW\_PUB.seq.\*  
2: /cgn2\_6/ptodata/1/pubpna/US06\_NEW\_PUB.seq.\*  
3: /cgn2\_6/ptodata/1/pubpna/US07\_NEW\_PUB.seq.\*  
4: /cgn2\_6/ptodata/1/pubpna/PCT\_NEW\_PUB.seq.\*  
5: /cgn2\_6/ptodata/1/pubpna/US03\_NEW\_PUB.seq.\*  
6: /cgn2\_6/ptodata/1/pubpna/US09\_NEW\_PUB.seq.\*  
7: /cgn2\_6/ptodata/1/pubpna/US10\_NEW\_PUB.seq.\*  
8: /cgn2\_6/ptodata/1/pubpna/US10\_NEW\_PUB.seq.\*  
9: /cgn2\_6/ptodata/1/pubpna/US11\_NEW\_PUB.seq.\*  
10: /cgn2\_6/ptodata/1/pubpna/US11\_NEW\_PUB.seq.\*  
11: /cgn2\_6/ptodata/1/pubpna/US11\_NEW\_PUB.seq.\*  
12: /cgn2\_6/ptodata/1/pubpna/US11\_NEW\_PUB.seq.\*  
13: /cgn2\_6/ptodata/1/pubpna/US60\_NEW\_PUB.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2424	100.0	1552	10	US-11-112-944-19
2	2424	100.0	1552	12	US-11-186-284-11
3	568	23.4	1701	7	US-10-063-703-73
4	568	23.4	1701	12	US-11-102-240-73

5	542.5	22.4	2771	8	US-10-276-233A-21	Sequence 21, Appl	
6	542.5	22.4	2775	9	US-11-072-175-87	Sequence 87, Appl	
7	409	16.9	1428	12	US-11-000-688-633	Sequence 633, App	
8	378	15.6	833	8	US-10-055-877-17	Sequence 17, Appl	
9	377	15.6	828	8	US-10-055-877-15	Sequence 15, Appl	
c	10	362	14.9	1224	12	US-11-136-527-1116	Sequence 1116, Ap
	11	362	14.9	1224	12	US-11-136-527-5212	Sequence 5212, Ap
12	359.5	14.8	1264	9	US-11-177-506-5	Sequence 5, Appli	
13	357.5	14.7	2357	12	US-11-000-688-845	Sequence 845, App	
14	342	14.1	1201	12	US-11-136-527-3817	Sequence 3817, Ap	
15	336	13.9	1858	12	US-11-136-527-556	Sequence 556, App	
16	333	13.7	1266	12	US-11-136-527-2459	Sequence 2459, Ap	
17	332	13.7	5787	12	US-11-169-041-59	Sequence 59, Appl	
18	312	12.9	7872	12	US-11-136-527-2235	Sequence 2235, Ap	
c	19	306	12.6	9858	7	US-10-893-483-185	Sequence 185, App
	20	306	12.6	36259	7	US-10-893-483-186	Sequence 186, App
21	304.5	12.6	3091	12	US-11-097-728-3	Sequence 3, Appli	
22	304.5	12.6	7941	12	US-11-097-728-1	Sequence 1, Appli	
23	304.5	12.6	8058	12	US-11-097-728-5	Sequence 5, Appli	
24	300	12.4	1104	12	US-11-000-688-914	Sequence 914, App	
25	300	12.4	1104	12	US-11-054-281-19	Sequence 19, Appl	
26	277	11.4	2531	8	US-10-131-826A-33	Sequence 33, Appl	
27	273	11.3	1670	8	US-10-131-826A-325	Sequence 325, App	
28	245	10.1	756	8	US-10-467-657-275	Sequence 275, App	
29	245	10.1	756	8	US-10-467-657-4007	Sequence 4007, Ap	
30	244	10.1	1400	12	US-11-136-527-4652	Sequence 4652, Ap	
c	31	224	9.2	629	12	US-11-136-527-3573	Sequence 3573, Ap
	32	198	8.2	1396	7	US-10-530-240-1	Sequence 1, Appli
33	197	8.1	600	12	US-11-136-527-7669	Sequence 7669, Ap	
34	179	7.4	1968	12	US-11-136-527-2510	Sequence 2510, Ap	
35	170.5	7.0	2363	12	US-11-049-348-3	Sequence 3, Appli	
36	149	6.1	600	12	US-11-136-527-7913	Sequence 7913, Ap	
37	148	6.1	2251	8	US-10-750-185-62901	Sequence 62901, A	
38	148	6.1	2251	8	US-10-750-623-62901	Sequence 62901, A	
39	144.5	6.0	2403	7	US-10-649-457-4	Sequence 4, Appli	
c	40	143.5	5.9	4990	12	US-11-128-061-932	Sequence 932, App
	41	143.5	5.9	4990	12	US-11-128-049-932	Sequence 932, App
42	143	5.9	1897	8	US-10-775-169-236	Sequence 236, App	
43	142	5.9	775	12	US-11-136-527-1683	Sequence 1683, Ap	
44	142	5.9	775	12	US-11-136-527-5779	Sequence 5779, Ap	
c	45	142	5.9	168516	12	US-11-121-086-3	Sequence 3, Appli

#### ALIGNMENTS

RESULT 1  
US-11-112-944-19  
; Sequence 19, Application US/11112944  
; Publication No. US20050244872A1  
; GENERAL INFORMATION:  
; APPLICANT: Harris, Cole  
; TITLE OF INVENTION: Breast Cancer Gene Expression Biomarkers  
; FILE REFERENCE: 05-325-US  
; CURRENT APPLICATION NUMBER: US/11/112,944  
; PRIOR FILING DATE: 2005-04-22  
; PRIOR APPLICATION NUMBER: US 60/564,757  
; PRIOR FILING DATE: 2004-04-23  
; NUMBER OF SEQ ID NOS: 29  
; SOFTWARE: PatentIn version 3.3  
; SEQ ID NO 19  
; LENGTH: 1552  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-11-112-944-19

Alignment Scores:  
Pred. No.: 7.51e-154  
Score: 2424.00  
Length: 1552  
Matches: 459  
Percent Similarity: 100.0%  
Best Local Similarity: 100.0%  
Query Match: 100.0%  
DB: 10  
Conservative: 0  
Mismatches: 0  
Indels: 0  
Gaps: 0





```
Db 231 CCTCAAAATCCAGTCCCTGCAACCCCTTCTCGGACACATATGTTGTTCTCCGCCCTCCT 290
Qy 127 AsnAsnAlaHisArgAspLys-GluGlyAspAspGlnSerHisTrpArgTyr----- 143
Db 291 GCTGGAGGTATTTGGATCTGGCTGCAGATGGGGTCAACACTGGACGTATGAGGGCCC 350
Qy 144 -GlyGlyAspProTrpProArgValSerProAlaCysAlaGlyArgPheGlnSerPr 163
Db 351 ACATGGTCAGACCATTTGGCCAGCCTCTTACCCTGAGTGTGAAACAATGCCAGTCGCC 410
Qy 163 oValAspIleArgProGlnLeuAlaPheCysProAlaLeuArgProLeuGluLeu 193
Db 411 CATCGATATTACAGACAGACAGTGTGACATTTGACCCCTGATTTGCCCTCTGCAGCCCCA 470
Qy 183 uGlyPheGlnLeuProLeuProGluLeuArgLeuArgAsnAsnGlyHisSerValGl 203
Db 471 CGGATATGACAGCCTGGCACCAGCCTTTGGACCTGCACAAACATGGCCACAGTGCA 530
Qy 203 nLeuThrLeuProProGlyLeuGluMetAlaLeuGlyProGlyArgGluTyrArgAlaLe 223
Db 531 ACTCTCTCTCCCTCTACCTCTGTATCTGGGT---GGACTTCCCGGAAATATGTAGTCGC 587
Qy 223 uGlnLeuHisLeuHisTrpGlyAlaAlaGlyArgPro--GlySerGluHisThrValGl 242
Db 588 CCAGCTCCACCTGCACCTGGGGTCAAAAGGATCCCCAGGGGGTGCAGAACCCAGATCAA 647
Qy 242 uGlyHisArgPheProAlaGluIleHisValHisLeu---SerThrAlaPheAlaAr 261
Db 648 CAGTGAAGCCACATTTGCAGAGCTCCACATTTGACATTTATGACTCTGATTCCTATGACAG 707
Qy 261 gValAspGluAlaLeuGlyArgProGlyGlyLeuAlaValLeuAlaPheLeuGluGl 281
Db 708 CTTGAGTGGGCTGCTGAGAGGCTCAGGGCCTGGCTGCTCGGCATCTCTAAATGAGGT 767
Qy 281 uGlyProGluLeuAsnSerAlaTrpGluGlnLeuLeuSerArgLeuGluIleAlaGl 301
Db 768 GGGTGAGACTTAAGAATATAGCTTTATGAACACTCTGTGAGTCACCTTGATGAAGTCAGGCA 827
Qy 301 uGluGlySerGluThrGlnValProGlyLeuAspIleSerAlaLeuLeuProSerAspPh 321
Db 828 TAAAGATCAGAGACCTCAGTGCTCTCCCTTCAACCTTAAGAGAGCTGCTCCCAACAGCT 887
Qy 321 eSerArgTyrPheGlnTrpGluGlySerLeuThrProProCysAlaGlnGlyValII 341
Db 888 GGGGCAGTACTTCGGCTACAATGGCTCGCTCACAACTCCCCCTTGCTTACCAGAGTGTGCT 947
Qy 341 eTrpThrValPheAsnGlnThrValMetLeuSerAlaLysGlnLeuHisThrLeuSerAs 361
Db 948 CTGGACAGTTTTATAGAAGGTCCACAGATTTCAATGGAAACAGCTGGAAGAAGCTTCAGGG 1007
Qy 361 pThrLeuTrpGlyProGlyAspSerArgLeuGlnLeu-----AsnPheArgAlaTh 378
Db 1008 GACATTTCTCCACAGAAGAGGAGCCCTTAAGCTTCTGGTACAGAACTACCGAGCCCT 1067
Qy 378 rGlnProLeuAsnGlyArgValIleGluAlaSerPheProAlaGlyValAspSerPr 398
Db 1068 TCAGCCTCTCAATCAGCGCATGGTCTTTGCTTCTTC----- 1104
Qy 398 oArgAlaAlaGluProValGlnLeuAsnSerCysLeuAlaAlaGlyAspIleLeuAlaLe 418
Db 1105 -----ATCCAAGCAGGATCCTCGTATACCACAGGTGAATGCTGAGTCT 1148
Qy 418 uValPheGlyLeuLeu-PheAlaValThrSerValAla-----PheLeuValGlnMetA 436
Db 1149 AGGTGTAGGAATCTTGGTTGGCTGTCTCGCCTTCCTCGCTCTTATTTCATTGCTAG 1208
Qy 436 rgArgGlnHisArgArgGlyThrIlys 444
Db 1209 AAAGATTCCGAAGAAGAGGCTGGAAA 1234
RESULT 4
US-11-102-240-73
```

```
; Sequence 73, Application US/11102240
; Publication No. US20050260647A1
; GENERAL INFORMATION:
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, Christopher J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Wood, William I.
; TITLE OF INVENTION: ANTIBODIES TO POLYPEPTIDES ENCODED BY A NUCLEIC ACID UNDEREXPRESSI
; FILE REFERENCE: P3230R1C106C
; CURRENT APPLICATION NUMBER: US/11/102,240
; PRIOR FILING DATE: 2005-04-08
; PRIOR APPLICATION NUMBER: 10/063662
; PRIOR FILING DATE: 2002-05-07
; PRIOR APPLICATION NUMBER: 10/006867
; PRIOR FILING DATE: 2001-12-06
; PRIOR APPLICATION NUMBER: PCT/US00/23328
; PRIOR FILING DATE: 2000-08-24
; PRIOR APPLICATION NUMBER: 60/170262
; PRIOR FILING DATE: 199-12-09
; NUMBER OF SEQ ID NOS: 170
; SEQ ID NO 73
; LENGTH: 1701
; TYPE: DNA
; ORGANISM: Homo Sapien
; FEATURE:
; NAME/KEY: unsure
; LOCATION: 1528
; OTHER INFORMATION: unknown base
US-11-102-240-73

Alignment Scores:
Pred. No.: 3.8e-29 Length: 1701
Score: 568.00 Matches: 151
Percent Similarity: 47.9% Conservative: 55
Best Local Similarity: 35.1% Mismatches: 176
Query Match: 23.4% Indels: 48
DB: 12 Gaps: 12

US-09-967-237B-2 (1-459) x US-11-102-240-73 (1-1701)
Qy 38 GlnArgLeuProArgMetGlnGluAspSerProLeuGlyGlySerSerGlyGluAsp 57
Db 24 GAGAGGGCAAGAGGAGGAGGAGAGATTTGTCCT-----GGGATCCAGAAACCCAT 74
Qy 58 AspProLeuGlyGluGluAspLeuProSerGluGluAspSerProArgGluGluAspPro 77
Db 75 GATACCTTACTGAACACCGAATCCCTTGGAAAGCCACAGACAGACAGCAAGAGAG 134
Qy 78 ProGlyGluGluAspLeuProGlyGluGluAspLeuProGlyGluGluAspLeuProGlu 97
Db 135 -----CAGAGATAAATACACTCACGCCAGGAGCTCGCTCTCTCTCTCTCTCTC 188
Qy 98 ValLysProLysSerGluGluGluGlySerLeuLysLeu-----GluAspLeu 113
Db 189 ACTCCTCC-----TCCCTCTCTGCGCTGCTAGTCTCTAGT 230
Qy 114 ProThrValGluAlaProGlyAspPro-----GlnGluProGln 126
Db 231 CCTCAAAATCCCAGTCCCTCGCACCCCTTCTCTGGGACACTATGTTGTTCTCCGCCCTCCT 290
Qy 127 AsnAsnAlaHisArgAspLys-GluGlyAspAspGlnSerHisTrpArgTyr----- 143
Db 291 GCTGGAGGTGATTTGGATCTGGCTGCAGATGGGGTCAACACTGGACGTATGAGGGCCC 350
Qy 144 -GlyGlyAspProProTrpProArgValSerProAlaCysAlaGlyArgPheGlnSerPr 163
Db 351 ACATGGTCAGACCATTTGGCCAGCCTTACCCCTGAGTGTGAAACAATGCCAGTCGCC 410
Qy 163 oValAspIleArgProGlnLeuAlaPheCysProAlaLeuArgProLeuGluLeu 183
Db 411 CATCGATATTACAGACAGACAGTGTGACATTTGACCCCTGATTTGCGCTGCTCTGCAGCCCCA 470
```

Qy	183	uglyPheGlnLeuProLeuProGluLeuArgAsnAsnGlyHisSerValGl	203
Db			
Db	471	CGGATATGACAGCCTGGCACCCAGGCCTTTGGACCTGCACAACAATGCCACACAGTGCA	530
Qy	203	nLeuThrLeuProProGlyLeuGluMetAlaLeuGlyProGlyArgGlnTyrArgAlaLe	223
Db			
Db	531	ACTCTCTCTGCCCTCTACCTCTGTATCTGGGT---GGACTTCCCCGAANAATATGTAGTCGC	587
Qy	223	uGlnLeuHisLeuHisTrpGlyAlaLaGlyArgPro---GlySerGluHisThrValGl	242
Db			
Db	588	CCAGCTCCACCTGCACTGGGGTCTCAGAAGGATCCCCAGGGGGGTTCAGAACACCAGATCAA	647
Qy	242	uGlyHisArgPheProAlaGluIleHisValHisLeu---SerThrAlaPheAlaAr	261
Db			
Db	648	CAGTGAAGCCACATTTCAGAGCTCCACATGTACATTATGACTCTCTGATTCTCTATGACAG	707
Qy	261	gValAspGluAlaLeuGlyArgProGlyGlyLeuAlaValLeuAlaAlaPheLeuGluGl	281
Db			
Db	708	CTTGAGTGAGGCTGCTCAGAGGCCCTCAGGGCTCGGCTGCTCGGGCATCCTAATTGAGGT	767
Qy	281	uGlyProGluGluAsnSerAlaTyrGluGlnLeuLeuSerArgLeuGluGluIleAlaGl	301
Db			
Db	768	GGGTGAGACTATGAAGAAATATAGCTTATGAACACATCTCGAGTCAC TTGCATGAAGTCAGGCA	827
Qy	301	uGluGlySerGlnThrGlnValProGlyLeuAspileSerAlaLeuLeuProSerAspPh	321
Db			
Db	828	TAAAGATCAGAGAGACCTCAGTGCCCTCCCTTCAACCTAAGAGAGCTGCTCCCCAACAGCT	887
Qy	321	eSerArgTyrPheGlnTyrGluGlySerLeuThrThrProProCysalaGlnGlyValIl	341
Db			
Db	888	GGGSCAGTACTTCGGCTACAAATGGCTCGCTCACAACCTCCCCTTGGCTACCAGAGTGTGCT	947
Qy	341	eTTPThrValPheAsnGlnThrValMetLeuSerAlaLysGlnLeuHisThrLeuSerAs	361
Db			
Db	948	CTGGACAGTTTTTTATAGAGGTCCTCAGATTCCAATGGACAGCTGGAAAAGCTTCAGGG	1007
Qy	361	pThrLeuTrpGlyProGlyAspSerArgLeuGlnLeu-----AsnPheArgAlaTh	378
Db			
Db	1008	GACATTTGTTCTCCACAGAGAGGAGGCCTCTAAGCTTCTGGTACAGAACTACCGAGGCCCT	1067
Qy	378	rGlnProLeuAsnGlyArgValIleGluAlaSerPheProAlaGlyValAspSerSerPr	398
Db			
Db	1068	TCAGCTCTCAATCAGGCCATGCTTGTGTTCTTTTC-----	1104
Qy	398	oArgAlaAlaGluProValGlnLeuAsnSerCysLeuAlaAlaGlyAspIleLeuAlaLe	418
Db			
Db	1105	-----ATCCAAGCAGCATCTCTGTATACCACAGGTCGAAATGCTGAGTCT	1148
Qy	418	uValPheGlyLeuLeu-PheAlaValThrSerValAla-----PheLeuValGlnMeta	436
Db			
Db	1149	AGGTGTAGGAATCTGTGTTGGCTGCTCTGCTCTCTCTCTGCTGTTTATTTCATGCTAG	1208
Qy	436	rArgGlnHisArgArgGlyThrlys	444
Db			
Db	1209	AAGATTTCGGAAGAAGAGGCTGGAAA	1234

## RESULT 5

US-10-276-233A-21
; Sequence 21, Application US/10276233A
; Publication No. US20050260572A1
; GENERAL INFORMATION:
; APPLICANT: DNA Chip Research Inc.
; APPLICANT: Hitachi Software Engineering Co., Ltd.
; TITLE OF INVENTION: A method of predicting cancer condition
; FILE REFERENCE: PH-1533-PCT
; CURRENT APPLICATION NUMBER: US/10/276,233A
; CURRENT FILING DATE: 2002-11-14
; PRIOR APPLICATION NUMBER: JP 2001-73063
; PRIOR FILING DATE: 2001-03-14
; PRIOR APPLICATION NUMBER: JP 2001-108503
; PRIOR FILING DATE: 2001-04-06
; PRIOR APPLICATION NUMBER: JP 2001-234807

  

290	GlyLeuAlaValLeuAlaAlaPheLeuGluGluGlyProGluGluAsnSerAlaTyrGlu	290
QY		
291	GlyLeuAlaValLeuAlaAlaPheLeuGluGluGlyProGluGluAsnSerAlaTyrGlu	291
QY		
611	GGCCTCGCTGTCTGGCTTTCTTCATTGAGATGGGC---TCCTTCAATCCGCTCATGAC	667
Db		
291	GlnLeuLeuSerArgLeuGluLuuIleAlaGluGluGlySerGluThrGlnValProGly	310
QY		
668	AGATCTTCAGTACCCTTCCAACATGTAAAGTACAAGGCCAGGAAGCATTTGCTCCCGGA	727
Db		
311	LeuAspIleSerAlaLeuLeuProSerAspPheSerArgTyrPheGlnTyrGluGlySer	330
QY		
728	TTCACACTTGAAGAGCTGCTTCGGGAGAGCACCGCTGAATATTACCGTACCGGGGGTCC	787
Db		
331	LeuThrThrProProCysAlaGlnGlyValIleTrpThrValPheAsnGlnThrValMet	350
QY		

Db 788 CTGACACACACCCCTTCCAAACCCCTGCTGCTCTGGACAGCTTTTCCGAACCCCGTGCAA 847  
Qy LeuSerAlaLysGlnLeuHisThrLeuSerAspThrLeuTrp-----Gly 365  
Db 848 ATTTCACAGGAGCAGCTGCTGGCTTTGGACAGACGCCCTGTACTGCACACATGGACGAC 907  
Qy ProGlyAspSerArgLeuGlnLeuAsnPheArgAlaThrGlnProLeuAsnGlyArgVal 385  
Db 908 CCTTCCCCAGAGAAATGATCAACAACCTCCGGCAGGTCAGAAAGTTTCGATGAGAGCTG 967  
Qy 386 IleGluAlaSerPheProAlaGlyValAspSerSerProArgAlaAlaGluProValGln 405  
Db 968 GTATACACCTCTCTCTCC-----CAA 988  
Qy 406 LeuAsnSerCysLeuAlaAlaGlyAspIleLeuAlaLeuValPheGlyLeuLeuPheAla 425  
Db 989 GTGCAAGTCTGTACTCGGCGAGCAGTGAAGTCTGGGCATCATCTCTCACTGGCCCTGGCT 1048  
Qy 426 -----ValThrSerValAlaPheLeuValGlnMetArgArgGln 438  
Db 1049 GGCATTCTTGGCATCTGTATTGTGGTGGTGTCTCCATTTTGGCTTTTTCAGAGGAAGAGT 1108  
Qy 439 HisArgArgGlyThrLysGlyValSerTyrArgProAlaGluValAlaGluThrGly 458  
Db 1109 ATCAAAAAGGTGATACAAAGGGAGTCAATTACAGCCAGCCACCAAGATGGAGACTGAG 1168  
Qy 459 Ala 459  
Db 1169 GCC 1171

## RESULT 6

US-11-072-175-87  
; Sequence 87, Application US/11072175  
; Publication No. US20060029944A1  
; GENERAL INFORMATION:  
; APPLICANT: Bristol-Myers Squibb Company  
; TITLE OF INVENTION: IDENTIFICATION OF GENES FOR PREDICTING ACTIVITY OF COMPOUNDS THAT  
; TITLE OF INVENTION: INTERACT WITH AND/OR MODULATE PROTEIN TYROSINE KINASES AND/OR  
; FILE OF INVENTION: PROTEIN TYROSINE KINASE PATHWAYS IN BREAST CELLS  
; FILE REFERENCE: D0273A CIP  
; CURRENT APPLICATION NUMBER: US/11/072,175  
; PRIOR FILING DATE: 2005-03-05  
; PRIOR APPLICATION NUMBER: US 60/406,385  
; PRIOR FILING DATE: 2002-08-27  
; PRIOR APPLICATION NUMBER: US 10/648,593  
; PRIOR FILING DATE: 2003-08-26  
; NUMBER OF SEQ ID NOS: 571  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 87  
; LENGTH: 2775  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-11-072-175-87

## Alignment Scores:

Pred. No.: 3.01e-27 Length: 2775  
Score: 542.50 Matches: 134  
Percent Similarity: 47.2% Conservative: 46  
Best Local Similarity: 35.2% Mismatches: 150  
Query Match: 22.4% Indels: 51  
DB: 9 Gaps: 11

US-09-967-237B-2 (1-459) x US-11-072-175-87 (1-2775)

Qy 114 ProThrValGluAlaProGlyAspProGlnAsnAlaHisArgAsp--- 132  
Db 77 CCGGCACAGCCGCGCCCGCCG---CCGACAGAGCCCGGAGATGCCCGCGCAGCCTG 133  
Qy 133 -----LysGluGlyAspAspGln----- 138  
Db 134 CACGCGCGCGCGTCTCTGCTGTATTAAGGACACAGCCTTCCAGCCCGGCCCA 193  
Qy 139 -----SerHisTrpArgTyr-----GlyGlyAspProProTrpProArgVal 152

## RESULT 7

US-11-000-688-633



```
; PRIOR APPLICATION NUMBER: 60/275,990
; PRIOR FILING DATE: 2001-03-14
; PRIOR APPLICATION NUMBER: 60/275,927
; PRIOR FILING DATE: 2001-03-14
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 512
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 17
; LENGTH: 833
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (21)..(806)
US-10-055-877-17

Alignment Scores:
Pred. No.: 1.18e-16 Length: 833
Score: 378.00 Matches: 98
Percent Similarity: 51.1% Conservative: 41
Best Local Similarity: 36.0% Mismatches: 116
Query Match: 15.6% Indels: 17
DB: 8 Gaps: 8

US-09-967-237B-2 (1-459) x US-10-055-877-17 (1-833)
Qy 131 ArgAspLysGluGlyAspAspGlnSerHis-TripArgTyr-----GlyGlyAspPr 147
Db 5 CGAGGCTCAGTGGGATGCGAGGCTCAGTGGGATACCGGAGCACAACGCTCTAT 64
Qy 147 oProTTPProArgValSerProAlaCysAlaGlyArgPheGlnSerProValAspLeAr 167
Db 65 TCACTGGAAGAAATTTTCCCTATTGTGTGTGTGATCAGCAATCTCCAATTCAGATTA 124
Qy 167 gProGlnLeuAlaAlaPheCysProAlaLeuArgProLeuGluLeuLeuGlyPheGlnLe 187
Db 125 AACCAAGAAGTGAATATGACTTCCCTCCGACCACTTAGTATC-----AAGTA 175
Qy 187 uProProLeuProGluLeuArgLeuArgAsnGlyHisSerValGlnLeuThrLeuPr 207
Db 176 TGACCAAGCTCAGTAAATCATCAGCAACAGCGGCCATCTCTCAATGTGACTTGA 235
Qy 207 oProGlyLeuGluMetAlaLeu-----GlyPro-----GlyArgGluTyrArgAlaLe 223
Db 236 TGACACAGACAACAAATCTCTCGTGGTGGTCTCTCACTGGAAGCTACAGGTTACG 295
Qy 223 uGlnLeuHisLeuHisTTPGlyAlaAlaGlyArgProGlySerGluHisThrValGluGl 243
Db 296 GCAGGTTCACTTCACTGGGGGTCGCTGTATGATGACCGGCTCCAGCACATAGTAGTGG 355
Qy 243 yHisArgPheProAlaGluLeuHisValValHisLeu---SerThrAlaPheAlaArgVa 262
Db 356 AGTGAGCTATGCTGCAGAGCTCCATGTGTCTACTGGAATTCAGACAAATACCCAGCTT 415
Qy 262 lAspGluAlaLeuGlyArgProGlyGlyLeuAlaValLeuAlaAlaPheLeuGluGl 282
Db 416 TGTGAGGAGCTCATGAACAGATGAGTGGCTGTCTGGAGTGTGTTTACAGATTGG 475
Qy 282 yProGluGluAsnSerAlaTyrGluGlnLeuLeuSerArgLeuGluGluLeuAlaGluGl 302
Db 476 T---GAACCTAATTCCTCCAAAGATTAATGACACTTGGATTCATTAAGAAA 532
Qy 302 uGlySerGluThrGlnValProGlyLeuAspIleSerAlaLeuLeuProSerAspPheSe 322
Db 533 GGGTAACAAACTCGATTCAAAATTTTGACCTATTGTCTCTGCTTCCACCATCCTGGGA 592
Qy 322 xArgTyrPheGlnTyrGluGlySerLeuThrProProCysAlaGlnGlyValIleTr 342
Db 593 C---TACTGACATATCTGCTCTCTTACAGTTCCACCTCTTCTTGAGAGTGTCAATG 649
Qy 342 pThrValPheAsnGlnThrValMetLeuSerAlaLysGlnLeu-----HisThrLe 359
Db 650 GATTGTTTAAAGCAACTATAACATCAGCTCTCAACAGCTGGCCAAATTCGAGTCT 709
```

```
Qy 359 userAspThrLeuTrpGlyProGlyAspSerArgLeuGlnLeuAsnPheArgAlaThrGl 379
Db 710 CCTGTGCACAGCGGAGGTGAAGCAGCAGCTTTTCTGTGAGCAATCAGCGCCACCACA 769
Qy 379 nProLeuAsnGlyArgValIleGluAlaSerPhe 390
Db 770 GCCTCTAAAGGCGCCAAAGTGAGAGCCTCTTTTC 803

RESULT 9
US-10-055-877-15
; Sequence 15, Application US/10055877
; Publication No. US20050288241A1
; GENERAL INFORMATION:
; APPLICANT: DeCristofaro, Marc
; APPLICANT: Padigaru, Muralidhara
; APPLICANT: Miller, Charles
; APPLICANT: Tchernev, Velizar
; APPLICANT: Zhong, Mei
; APPLICANT: Anderson, David
; APPLICANT: Ballinger, Robert
; APPLICANT: Gerlach, Valerie
; APPLICANT: Spytek, Kimberly
; APPLICANT: Ratelli, Luca
; APPLICANT: Kekuda, Ramesh
; APPLICANT: Guo, Xiaojia
; APPLICANT: Zerhusen, Bryan
; APPLICANT: Andrew, David
; APPLICANT: Mezes, Peter
; APPLICANT: Patturajan, Meera
; APPLICANT: Burgess, Catherine
; APPLICANT: Eisen, Andrew
; APPLICANT: Wolenc, Adam
; APPLICANT: Baumgartner, Jason
; APPLICANT: Shimkets, Richard
; APPLICANT: Gusev, Vladimir
; APPLICANT: Verniet, Corine
; APPLICANT: Taupier Jr., Raymond
; APPLICANT: Pena, Carol
; APPLICANT: Shenoy, Suresh
; APPLICANT: Li, Li
; APPLICANT: Casman, Stacie
; APPLICANT: Boldog, Ference
; TITLE OF INVENTION: Novel Polypeptides and Nucleic Acids Encoded Thereby
; FILE REFERENCE: 21402-251
; CURRENT APPLICATION NUMBER: US/10/055,877
; PRIOR FILING DATE: 2002-01-22
; PRIOR APPLICATION NUMBER: 60/262,892
; PRIOR FILING DATE: 2001-01-19
; PRIOR APPLICATION NUMBER: 60/263,598
; PRIOR FILING DATE: 2001-01-23
; PRIOR APPLICATION NUMBER: 60/263,799
; PRIOR FILING DATE: 2001-01-24
; PRIOR APPLICATION NUMBER: 60/264,117
; PRIOR FILING DATE: 2001-01-25
; PRIOR APPLICATION NUMBER: 60/264,139
; PRIOR FILING DATE: 2001-01-25
; PRIOR APPLICATION NUMBER: 60/264,478
; PRIOR FILING DATE: 2001-01-26
; PRIOR APPLICATION NUMBER: 60/263,351
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: 60/272,870
; PRIOR FILING DATE: 2001-03-02
; PRIOR APPLICATION NUMBER: 60/275,990
; PRIOR FILING DATE: 2001-03-14
; PRIOR APPLICATION NUMBER: 60/275,927
; PRIOR FILING DATE: 2001-03-14
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 512
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 15
; LENGTH: 828
; TYPE: DNA
```

ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: CDS  
LOCATION: (16)..(801)  
US-10-055-877-15

## Alignment Scores:

Pred. No.: 1,37e-16 Length: 828  
Score: 377.00 Matches: 95  
Percent Similarity: 51.3% Conservative: 39  
Best Local Similarity: 36.4% Mismatches: 111  
Query Match: 15.6% Indels: 16  
DB: 8 Gaps: 8

US-09-967-237B-2 (1-459) x US-10-055-877-15 (1-828)

```
Qy 141 TrpArgTyr-----GlyGlyAspProProTrpProArgValSerProAlaCysAla 157
Db 31 TGGGGATACCGGAGCACAACGGTCTATTCACTGGAAGGAATTTTCCCTATTGCTGAT 90

Qy 158 GlyArgPheGlnSerProValAspIleArgProGlnLeuAlaAlaPheCysProAlaLeu 177
Db 91 GGTGATCAGCAATCTCCAATTGAGATTAAACCAAGAGTGAATATGACTCTTCCTC 150

Qy 178 ArgProLeuGluLeuGlyPheGlnLeuProProLeuProGluLeuArgLeuArgAsn 197
Db 151 CGACCATAGTATC-----AAGTATGACCAAGCTCAGCTAAATCATCAGCAAC 201

Qy 198 AsnGlyHisSerValGlnLeuThrLeuProProGlyLeuGluMetAlaLeu----- 214
Db 202 AGCGCCATCTCTCAATGTTGACTTTGATGACACAGACAACAAATCAGTTCTCGGTGT 261

Qy 215 GlyPro---GlyArgGluTyrArgAlaLeuGlnLeuHisLeuHisTrpGlyAlaAlaGly 233
Db 262 GGTCTCTCTCACTGGAAGCTACAGGTACAGGTACCGGAGTTTCACTTCACTGGGGTCCGCTGAT 321

Qy 234 ArgProGlySerGluHisThrValGluGlyHisArgPheProAlaGluIleHisValVal 253
Db 322 GACCAGGCTCCGAGCACATAGTAGAGTGTAGTATGCTGCAGAGCTCCATGTTGTT 381

Qy 254 HisLeu---SerThrAlaPheAlaArgValAspGluAlaLeuGlyArgProGlyGlyLeu 272
Db 382 CACTGGAATTCAGACAATAACCCAGCTTTTGTGAGGAGCTCATGAACACAGATGGACTG 441

Qy 273 AlaValLeuAlaAlaPheLeuGluGluGlyProGluGluAsnSerAlaTyrGluGlnLeu 292
Db 442 GCTGCTCTGGGAGTGTGTTTACAGGTGGT---GAACCTAATCCCACTGCAAAAGATT 498

Qy 293 LeuSerArgLeuGluGluIleAlaGluGlySerGluThrGlnValProGlyLeuAsp 312
Db 499 ACTGACACTTTGGATTCCATTAAAGAAAGGGTAAACAACTCGATTCAAAATTTTGAC 558

Qy 313 IleSerAlaLeuLeuProSerAspPheSerArgTyrPheGlnTyrGluGlySerLeuThr 332
Db 559 CTATTGCTCTGCTTCCACCCTCTGGGAC---TACTGGACATATCTCGGTCTCTTTACA 615

Qy 333 ThrProProCysAlaGlnGlyValIleTrpThrValPheAsnGlnThrValMetLeuSer 352
Db 616 GTTCCACCTCTTCTTGAGAGTGTACATGGATTGTTTAAAGCAACCTATAAATCAGC 675

Qy 353 AlaLysGlnLeu-----HisThrLeuSerAspThrLeuTyrProGlyAspSer 369
Db 676 TCTCAACAGCTGGCCAAATTTCCGAGCTCTCTGTGACAGCGGAGGTGAAGCAGCAGCT 735

Qy 370 ArgLeuGlnLeuAsnPheArgAlaThrGlnProLeuAsnGlyArgValIleGluAlaSer 389
Db 736 TTTCTGTGTGAGCAATCACCGGCCACCAAGCTCTTAAAGGGCCCAAGTGAAGAGCTCT 795

Qy 390 Phe 390
Db 796 TTC 798
```

RESULT 10

US-11-136-527-1116/c  
Sequence 1116, Application US/11136527  
Publication No. US20050287570A1  
GENERAL INFORMATION:

APPLICANT: Wyeth  
APPLICANT: Mounts, William M  
TITLE OF INVENTION: Probe Arrays For Expression Profiling of Rat Genes  
FILE REFERENCE: 031896-041000 (AM101086)  
CURRENT APPLICATION NUMBER: US/11/136,527  
CURRENT FILING DATE: 2005-05-25  
PRIOR APPLICATION NUMBER: US 60/574,294  
PRIOR FILING DATE: 2005-05-26  
NUMBER OF SEQ ID NOS: 362830  
SOFTWARE: PatentIn version 3.2  
SEQ ID NO 1116  
LENGTH: 1224  
TYPE: DNA  
ORGANISM: Rattus norvegicus  
US-11-136-527-1116

## Alignment Scores:

Pred. No.: 1,96e-15 Length: 1224  
Score: 362.00 Matches: 105  
Percent Similarity: 49.2% Conservative: 56  
Best Local Similarity: 32.1% Mismatches: 112  
Query Match: 14.9% Indels: 54  
DB: 11 Gaps: 11

US-09-967-237B-2 (1-459) x US-11-136-527-1116 (1-1224)

```
Qy 166 IleArgProGlnLeuAlaAlaPheCysProAlaLeuArgProLeuGluLeuLeuGlyPhe 185
Db 1200 ATAGACCC---CTCGTGTGTGGATGTCGCTGTCTCC-----AAATTACGTGTGTG 1150

Qy 186 GlnLeuProLeuProGluLeu-ArgLeuArgAsnAsnGlyHisSerValGlnLeuThr 205
Db 1149 CAG-----GGACTGTGAAGTCCACCAACGATGGACACACCATTCAGTCAT 1105

Qy 205 rLeuProProGlyLeuGluMetAlaLeuGlyPro-----GlyArgGluTyrArgAl 222
Db 1104 CTTGAAGTCGAATCAGTTTGTTCAGGAGGACCCCTGCTCCTCAGGAGCAGGAGTTGAAC 1045

Qy 222 aLeuGlnLeuHisLeuHisTrpGlyAlaAlaGlyArgProGlySerGluHisThrValG 242
Db 1044 GTATGAAGTTAGATTTTCACTGGGGAAGAGAAACACCGCTGTTCTGACGACACGGTCAA 985

Qy 242 uGlyHisArgPheProAlaGluIleHisValHisLeu---SerThrAlaPheAlaAr 261
Db 984 TTCAAAGCCTTCCCATGGAGCTCCACCTGATCCACTGGAATTCACGCTGTTTGGCAG 925

Qy 261 gValAspGluAlaLeuGlyArgProGlyGlyLeuAlaValLeuAlaAlaPheLeuGlu 281
Db 924 CATCGATGAGCGGTGGGAAACMCSASGGCATTTGTCATCATTCGACTGTTTCGCCAGAT 865

Qy 281 uGlyProGluGluAsnSerAlaTyrGluGlnLeuLeuSerArgLeuGluIleAlaG 301
Db 864 AGGG---AAGGAGCATTTGGCTTGAAGGCTGTGACTGAGATACTTCAGGATATCCAATA 808

Qy 301 uGluGlySerGluThrGlnValProGlyLeuAspIleSerAlaLeuLeuProSerAspPh 321
Db 807 CAAGGGAAAAATCCAAAACAATTCATGCTTTAATCTTAACACTTTATTACCAACCCCTCT 748

Qy 321 eSerArg---TyrPheGlnTyrGluGlySerLeuThrThrProCysAlaGlnGlyVa 340
Db 747 TCTGCGGATTTACTGGGTCTATGAAGGATCTCTTACTATTTCACCTTGAGTGAAGGAGT 688

Qy 340 lIleTrpThrValPheAsnGlnThrValMetLeuSerAlaLysGlnLeu----- 356
Db 687 TACCTGGATATTATCCGATACCCATTATATATCCAGCTGCAGATAGAAGAAATTCG 628

Qy 357 -----HisThrLeuSerAspThrLeuTyrGlyProGlyAspSerArgLeu 372
Db 627 AAGACTGAGGACACATGTTTAAAGGGGCGAGAACTGTGCAAGGGCTGTGACGGGATTTTGGG 568
```

```
Qy 372 nLeuAsnPheArgAlaThrGlnProLeuAenGlyArgValIleGluAlaSerPhe----- 390
Db 567 AGATAATTTCCGACCTACCCAGCCCTTGAGTGACAGAGTATCCGAGCAGCAATTTTCAGTA 508
Qy 391 -ProAlaGlyValAspSerPro----- 398
Db 507 GCCAGAGAAACCTGAAACAAGCCCATTTGCATCAGGGAAGAGACGTTGTCYCACAGAG 448
Qy 399 -----ArgAlaAlaGlu-ProValGlnLeuAsnSerCysAl 410
Db 447 TCCTTTGATCAGGAGTGGAAACTCTGACGCGGAGCTTCCTGTTAACTTCAAGCCCTGCA 388
Qy 410 euAla-----AlaGlyAspIleLeuAlaLeu-ValPheGlyLeuLeu 423
Db 387 TTATCTTCAGTTACTCGGCTTTGATGGAGATCTGTGRCATTTGCTGTATACATGATGTT 328
Qy 424 PheAlaValThrSerValAla-PheLeuValGlnMetArgArgGlnHisArgGlyTh 443
Db 327 TATGAGATATTAGAAAGTGGCTGTTCACTATAAAGAAATCAGTGTGTCATGCACACMCCAC 268
Qy 443 rLysGlyGlyVal 447
Db 267 AGCTTCTGGAATT 255

RESULT 11
US-11-136-527-5212
; Sequence 5212, Application US/11136527
; Publication No. US20050287570A1
; GENERAL INFORMATION:
; APPLICANT: Wyeth
; APPLICANT: Mounts, William M
; TITLE OF INVENTION: Probe Arrays For Expression Profiling of Rat Genes
; FILE REFERENCE: 031896-041000 (AM101086)
; CURRENT APPLICATION NUMBER: US/11/136,527
; CURRENT FILING DATE: 2005-05-25
; PRIOR FILING DATE: 2005-05-26
; NUMBER OF SEQ ID NOS: 362830
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 5212
; LENGTH: 1224
; TYPE: DNA
; ORGANISM: Rattus norvegicus
US-11-136-527-5212

Alignment Scores:
Pred. No.: 1.96e-15 Length: 1224
Score: 362.00 Matches: 105
Percent Similarity: 49.2% Conservative: 56
Best Local Similarity: 32.1% Mismatches: 112
Query Match: 14.9% Indels: 54
DB: 12 Gaps: 11

US-09-967-237B-2 (1-459) x US-11-136-527-5212 (1-1224)

Qy 166 IleArgProGlnLeuAlaAlaPheCysProAlaLeuArgProLeuGluLeuLeuGlyPhe 185
Db 25 ATACGACCC---CTCGCTGTGGATGTCGCTGTCTCC-----AAATTACGTGGTGTG 75
Qy 186 GlnLeuProProLeuProGluLeu-ArgLeuArgAsnAsnGlyHisSerValGlnLeuTh 205
Db 76 CAG-----GGAGTGTGAAGTCCACCATGATGGACACACCATTCAGTGCAT 120
Qy 205 rLeuProProGlyLeuGluMetAlaLeuGlyPro-----GlyArgGluTyArgAl 222
Db 121 CCTGAAGTCGAATCAGTTTGTTCAGGAGACCCCTGCTCCCTCAGGACGAGGAGTTGAACT 180
Qy 222 aLeuGlnLeuHisLeuHisThrGlyAlaAlaGlyArgProGlySerGlySerGlyHisThrValG 242
Db 181 GTATGAGTTAGATTTCACCTGGGGAAGAGAAACACGCGTGGTCTCTGAGCACACGGTCAA 240
Qy 242 uGlyHisArgPheProAlaGluIleHisValHisLeu---SerThrAlaPheAlaAr 261
```

```
Db 241 TTTCAAAGCCTTCCCATGAGCTCCACCTGAAATTCACGCTGTTGGCAG 300
Qy 261 gValAspGluAlaLeuGlyArgProGlyGlyLeuAlaValLeuAlaAlaPheLeuGluG 281
Db 301 CATCGATAGCGGTTGGGAAACMCASAGCATTTGCATCATTTGCATGTTCTGTCAGAT 360
Qy 281 uGlyProGluGluAsnSerAlaTyrgluGlnLeuLeuSerArgLeuGluIleAlaG 301
Db 361 AGGG---AAGGAGCATGTTGGCTTGAAGCGTGTGACTGAGATATTCAGGATATCAATA 417
Qy 301 uGluGlySerGluThrGlnValProGlyLeuAspIleSerAlaLeuLeuProSerAspPh 321
Db 418 CAAGGAAAAATCCAAACAAATTCATGCTTTATCTAAACACTTTATTTACAGACCTCT 477
Qy 321 eSerArg---TyPheGlnTyrgluGlySerLeuThrProProCysAlaGlnGlyVa 340
Db 478 TCTGCGGGATTACTGGGTCTATGAGAGATCTCTTACTATTTCCACCTTGCAGTGAAGGAGT 537
Qy 340 lIleThrThrValPheAsnGlnThrValMetLeuSerAlaLysGlnLeu----- 356
Db 538 TACCTGGATATTATCCGATACCATTAACTATATCCAGCTGCAGATAGAGAATTTTCG 597
Qy 357 -----HisThrLeuSerAspThrLeuTyrglyProGlyAspSerArgLeuG 372
Db 598 AAGACTGAGGACACATGTTAAGGGGCGACACTYGTGCGAAGGCTGTGACGGGATTTTGGG 657
Qy 372 nLeuAsnPheArgAlaThrGlnProLeuAsnGlyArgValIleGluAlaSerPhe----- 390
Db 658 AGATAATTTCCGACCTACCCAGCCCTTGAGTGACAGAGTATCCGAGCAGCAATTTTCAGTA 717
Qy 391 -ProAlaGlyValAspSerPro----- 398
Db 718 GCCAGAGAAACCTGAAACAAGCCCATYTGTCATCAGGGAAGAGACGTTGGTCTYCACAGAG 777
Qy 399 -----ArgAlaAlaGlu-ProValGlnLeuAsnSerCysAl 410
Db 778 TCCTTTGATGAGGAGTGGAAACTCTTWAGCGCGGAGCTTCTGTGTAACTTCAAGCCCTGCA 837
Qy 410 euAla-----AlaGlyAspIleLeuAlaLeu-ValPheGlyLeuLeu 423
Db 838 TTATCTTCAGTTACTCGGCTTTGATGGAGATCTGTGRCATTTGCTGTATACATGATGTT 897
Qy 424 PheAlaValThrSerValAla-PheLeuValGlnMetArgArgGlnHisArgGlyTh 443
Db 898 TATGAGATATTAGAGTGGCTGTTCACTATAAAGAAATCAGTGTGTCATGCACACMCCAC 957
Qy 443 rLysGlyGlyVal 447
Db 958 AGCTTCTGGAATT 970

RESULT 12
US-11-177-506-5
; Sequence 5, Application US/11177506
; Publication No. US20060029956A1
; GENERAL INFORMATION:
; APPLICANT: Beyer, Wayne F.
; APPLICANT: Venetta, Thomas M.
; APPLICANT: Groelke, John W.
; APPLICANT: Blaesius, Rainer H.
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE
; FILE REFERENCE: 46143/294851
; CURRENT APPLICATION NUMBER: US/11/177,506
; PRIOR FILING DATE: 2005-07-08
; PRIOR FILING DATE: 2004-07-09
; NUMBER OF SEQ ID NOS: 52
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 1264
; TYPE: DNA
; ORGANISM: Homo sapiens
```

[illegible]

```
Qy 144 -GlyGlyAspProProTrpProArgValSerProAlaCyAAlaGlyArgPheGlnSerPr 163
Db 713 CAACGGCTCTGACCACTGGATGAACCTTTTCCAAATGCAAGGGGAAACCAAGTCGCC 772
Qy 163 oValAspIleArgProGlnLeuAlaAlaPheCyProAlaLeuArgProLeuGluLeuLe 183
Db 773 CGTTGAGCTCATACTAAGACATCAGCATGACCTTCTCTGAGCCATCGTGTGTGG-- 830
Qy 183 uGlyPheGlnLeuProProLeuProGluLeuArgLeuArgAsnAsnGlyHisSerValG1 203
Db 831 -----TCATTATGATGGTGGCTCTGCGCAAGACCATCTCTGAATAATGGGAAGACCTGCCG 883
Qy 203 nLeuThrLeuProProGlyLeuGluMetAlaLeu-----GlyPr 216
Db 884 AGTTGTATTTGATGATATCTATGATAGTCAATGCTGAGAGGGGGTCTCTCCCTGGACC 943
Qy 216 oGlyArgGluTyArgAlaLeuGlnLeuHisHisTrpGlyValAlaAlaGlyArgProG1 236
Db 944 C-----TACGACTTCGCCAGTTTCACTTCCTGAGGGGCTCTTCGGATGATCATGG 994
Qy 236 ySerGluHisThrValGluGlyHisArgPheProAlaGluLeuHisValHisLeuSe 256
Db 995 CTCTGAGCACACCGTGGATGAGTCAAGTATGCGAGCGGAGCTTCATTTGTTCTCACTGGAA 1054
Qy 256 rThrAlaPheAlaArgValAspGluAlaLeuGlyArgProGlyGlyLeuAlaValLeuAl 276
Db 1055 CCGAAGTATAACACTTTTAAAGAAAGCCCTGAAAGCAGCGGATGGGATCGCTGTGATTGG 1114
Qy 276 aAlaPheLeuGluGluGlyProGluGluAanSerAlaTyArgGlnLeuLeuSerArgLe 296
Db 1115 CATTTTCTGAAGATAGGA---CATGAGATGGGAGTTCAGATTTTCTTGTATGCATT 1171
Qy 296 uGluGluAlaGluGlySerGluThrGlnValProGlyLeuAspIleSerAlaLe 316
Db 1172 GGACAAGATTAAAGACAAAGGCGAAGGAGCGCCCTTCCAAAGTTTGACCCATCTGCCT 1231
Qy 316 uLeuProSerAspPheSerArgTyPheGlnTyArgGlySerLeuThrThrProProCy 336
Db 1232 GTTCCCGGCA---TGCCGGGACTACTGAGACTTACCAGGGCTCATTCACACGCGCGCCGTG 1288
Qy 336 sAlaGlnGlyValIleTrpThrValPheAsnGlnThrValMetLeuSerAlaLysGlnLe 356
Db 1289 CGAGGAATGATGTGTGGCTGTCTGCTGAAGAGGCCATACCGTGAGCTCTGACCATGAT 1348
Qy 356 uHisThrLeuSerAspThrLeuTrpGlyProGlyAspSer-----ArgLeuGlnLe 373
Db 1349 GGCCAAGCTGGGAGCCCTCTCTCCAGTGTGTGAACAGCAGGCCGCCAGTGCCTCTTGTGAG 1408
Qy 373 uAsnPheArgAlaThrGlnProLeuAsnGlyArgValIleGluAlaSerPhe----- 390
Db 1409 CAATGGCGACCTTCCACAGCTTATCAATAACAGGGTGTGTGAGAGCTTCTTCAATGAGG 1468
Qy 391 -----ProAlaGlyValAspSer----- 396
Db 1469 CTGCTGGATCTTGGCCCTTTCAGGAAGGAACCTACCATTTGGAGAGCTTGGTTCCTTGC 1528
Qy 396 ----- 396
Db 1529 CTCCTCTGTGTCTTACTTCCAGTCTATTTCATTTTCCACACTGAGCAATGAATGTG 1588
Qy 397 -----SerProArgAlaAlaGluProValGlnLeuAsnSerCysLeuAlaAlaG1 413
Db 1589 AGAGATGTGTGTCACCAAGATCTAAGTTACTTGTGTGAAGAAAGTACTTTTCGACAAGATC 1648
Qy 413 yAspIleLeuAla-----LeuValPheGlyLeuLeuPheAlaValThrSerVa 429
Db 1649 TAATATGAAGCATAGATTTCACATTGTGATCTCTGTAATATCATCTTTCTTATAAAGT 1708
Qy 429 lAlaPheLeuValGlnMetArgArgGlnHisArgArg 441
Db 1709 AGCATTTTGTGAAGTTTCAAGAAAGAAACAGA 1745
```

```
US-11-136-527-3817
; Sequence 3817, Application US/11136527
; Publication No. US20050287570A1
; GENERAL INFORMATION:
; APPLICANT: Wyeth
; APPLICANT: Mounts, William M
; TITLE OF INVENTION: Probe Arrays For Expression Profiling of Rat Genes
; FILE REFERENCE: 031896-041000 (AM101086)
; CURRENT APPLICATION NUMBER: US/11/136,527
; PRIOR FILING DATE: 2005-05-25
; PRIOR APPLICATION NUMBER: US 60/574,294
; PRIOR FILING DATE: 2005-05-26
; NUMBER OF SEQ ID NOS: 362830
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 3817
; LENGTH: 1201
; TYPE: DNA
; ORGANISM: Rattus norvegicus
US-11-136-527-3817

Alignment Scores:
Pred. No.: 4.25e-14 Length: 1201
Score: 342.00 Matches: 93
Percent Similarity: 48.2% Conservative: 38
Best Local Similarity: 34.2% Mismatches: 113
Query Match: 14.1% Indels: 28
DB: 12 Gaps: 10

US-09-967-237B-2 (1-459) x US-11-136-527-3817 (1-1201)

Qy 140 HisTrpArgTyR-----GlyGlyAspProPro 148
Db 106 CACTGGCTTCCAGCATCTTACGAGAGAAACACAGCAACTGTGCGCGCCCTCTC 165
Qy 149 Trp-----ProArgValSerProAlaCysAlaGlyArgPheGlnSerProValAspIle 166
Db 166 TGACCTGGCCCATGTCTCTATGACCCGAGCTGGCAGCTCCAGGCTC-----TCC 219
Qy 167 ArgProGlnLeuAlaAlaPheCysProAlaLeuArgProLeuGluLeuGlyPheGln 186
Db 220 CAGTGGCAGGATAGTGTCTATGACCCGAGCTGGCAGCTC-----TCC 270
Qy 187 LeuProProLeuProGluLeuArgLeuArgAsnAsnGlyHisSerValGlnLeuThrLeu 206
Db 271 TATGATGTGCTGCTCTGAGATACCTCTGGAACACTGGTTACTTCTTCCAGGTGGAGTTT 330
Qy 207 ProProGlyLeuGlu-----MetAlaLeuGlyPro---GlyArgGluTyArgAla 222
Db 331 GACGATTCCTGTGAGGAGTCAGGGATCAGTGGTGGGCTCTGGGAAACCACTACAGGCTG 390
Qy 223 LeuGlnLeuHisLeuHisTrpGlyAlaAlaGlyArgProGlySerGlnHisThrValGlu 242
Db 391 AAGCAGTTTCACTTCCACTGGGAGCAACAGATGAATGGGGCTCTGAGCACATGGTGGAC 450
Qy 243 GlyHisArgPheProAlaGluIleHisValValHisLeu---SerThrAlaPheAlaArg 261
Db 451 GGCCATGTCTTACCCGCTGAGCTCCATTTGGTTCTACTGGAAATTCATGAATATGAANAAT 510
Qy 262 ValAspGluAlaLeuGlyArgProGlyGlyLeuAlaValLeuAlaAlaPheLeuGluGlu 281
Db 511 TACAAGAAAGCCACACCGGGGAGAGATGGATCGCGGTGATTGGAGTGTCTTCTGAAGCTC 570
Qy 282 GlyProGluGluAsnSerAlaTyArgGlnLeuLeuSerArgLeuGluGluIleAlaGlu 301
Db 571 GGG---GCCCATCAGGAGCCCTGACAGAGGCTGTGGACATCTTCCGCGGAAGTAAGACAC 627
Qy 302 GluGlySerGluThrGlnValProGlyLeuAspIleSerAlaLeuLeuProSerAspPhe 321
Db 628 AAGAACACACAGCATGATGCATGCGGGCCCTTTGACCCCTTCTTGTCTGTCTGTCTGCGC 684
Qy 322 SerArgTyRPhenTyArgGlySerLeuThrProProCysAlaGlnGlyValIle 341
Db 685 CGGGATTACTGGACCTACCTGGCTCCCTCACCACCCCACTGCTGCTGAGTCAGTCAGTCACC 744
```



**THIS PAGE BLANK (USPTO)**